CombiGlide 1.0

User Manual

Schrödinger Press

Revision A, April 2006

Contents

Document Conventions

In addition to the use of italics for names of documents, the font conventions that are used in this document are summarized in the table below.

Table 3.1.

Font	Example	Use
Sans serif	Project Table	Names of GUI features, such as panels, menus, menu items, buttons, and labels
Monospace	`$SCHRODINGER/maestro`	File names, directory names, commands, environment variables, and screen output
Italic	*filename*	Text that the user must replace with a value
Sans serif uppercase	CTRL+H	Keyboard keys

In descriptions of command syntax, the following UNIX conventions are used: braces `{ }` enclose a choice of required items, square brackets `[ ]` enclose optional items, and the bar symbol | separates items in a list from which one item must be chosen. Lines of command syntax that wrap should be interpreted as a single command.

In this document, to *type* text means to type the required text in the specified location, and to *enter* text means to type the required text, then press the ENTER key.

References to literature sources are given in square brackets, like this: [10].

Introduction to CombiGlide

Combinatorial chemistry and parallel synthesis techniques allow for the generation of orders of magnitude more compounds than can in reality be prepared or screened. Thus, there is a need for the development of methods for the rational selection of the optimal members of potential libraries for synthesis. In recent years, combinatorial library design has shifted toward small focused libraries that are biased toward a specific target or class of targets and exhibit optimal drug-like physiochemical properties. Structure-based design strategies can be implemented in the design of focused libraries when a 3D structure of the target is available.

CombiGlide is a computational tool for the rapid virtual screening of combinatorial libraries to eliminate unpromising compounds. CombiGlide uses technology from the extra precision (XP) mode of our docking product, Glide, to dock molecules in which side-chains are substituted onto a defined core structure. The centerpiece of the approach is a proprietary selection protocol that rapidly determines from the docking results which members of the virtual library have the highest likelihood of binding well to the target. These library members are then enumerated and docked. A variety of library selection strategies is provided, including methods for the incorporation of predicted ADME properties into the overall decision process. CombiGlide can be used in the lead discovery or lead optimization phases of a program. In the lead discovery phase, various cores, sites of diversity, and chemistries can be evaluated. In the lead optimization phase, the importance of functionality at different positions on the initial hit can be evaluated along with the screening of extensive collections of side chains at each position.

1.1 CombiGlide Features

CombiGlide includes the following capabilities and features:

- Performs rapid screening of large virtual combinatorial libraries against any target for which a 3D model is available
- Is orders of magnitude faster than docking the entire library
- Uses a “core plus side chains” approach
- Provides automated reagent file preparation: 2D to 3D conversion, generation of reasonable ionization and tautomeric states, stereoexpansion, assignment of attachment points
- Offers extensive flexibility in initial core placement

- Performs flexible docking using the extra precision (XP) mode of Glide
- Uses an intuitive, user-friendly wizard-based GUI for setting up and monitoring jobs and for visualization of docked poses
- Provides multiple post-docking library selection strategies and options
- Allows for incorporation of predicted ADME properties into selection process
- Enumerates complete combinatorial libraries on request

1.2 Overview of the Focused Library Design Process

The main intended use of CombiGlide is for the design of a focused combinatorial library. The steps in this process are summarized here. The last five steps constitute the combinatorial screening process (which is summarized with some more detail in Chapter 5).

1. Determine the desired synthetic approach.

   The first step is to identify the chemistry associated with the desired library: the sequence of reactions that leads to the desired product. This step is discussed in more detail in Chapter 3.

2. Create or obtain reagent files.

   When the synthetic approach is determined, you must then select an appropriate set of reagents, and obtain files containing the structures of the reagents. This step is discussed in more detail in Chapter 3.

3. Identify and prepare the core.

   The core is the structural element that is constant throughout the library. The side chains are built onto the core using the reagents in the reagent files. You must supply a molecule that defines the core and ensure that it is a 3D, minimized structure, using LigPrep or MacroModel, for example. This step is discussed in more detail in Chapter 3.

4. Prepare receptor and receptor grids.

   The receptor must be prepared for the docking stage, and the receptor grids for the docking must be generated. These tasks are performed in the Protein Preparation and Receptor Grid Generation panels, and are described in brief in Chapter 4 and in detail in the *Glide User Manual*.

5. Prepare reagent files.

   The structures in the reagent files must be converted to 3D, all-atom structures and minimized. The bonds that will be replaced when the side chains are attached to the core must

also be defined. These tasks are performed in the Reagent Preparation panel, and are described in detail in Chapter 4.

6. Define attachment points and associated reagent files.

   After preparing the reagent files and the core, you define the points at which the side chains from the reagents will be attached to the core to construct the library, and associate a reagent file with each attachment point. These tasks are performed in the Define Combinations step of the Combinatorial Screening panel. This step is described in detail in Chapter 6.

7. Select parameters for the docking runs.

   Structures are docked using CombiGlide XP docking, which makes use of Glide XP technology with some variations. The appropriate Glide options for docking can be set in the Configure Docking step of the Combinatorial Screening panel. This step is described in detail in Chapter 7.

8. Define the core poses.

   A set of core poses is obtained by docking selected molecules. These core poses are used as initial positions for docking the library members. In the Define Core Poses step of the Combinatorial Screening panel you can choose which molecules to dock for the core poses (they can be different from the one you selected above) and set constraints on the core position. This step is described in detail in Chapter 7.

9. Dock the library members.

   The docking step is done in three stages. First, the core-containing molecules are docked to generate a set of core poses. Next, structures resulting from a single substitution at any position on the core structure are docked. The unpromising reagents at each position are screened out, and the reduced list is used to dock the best of the fully substituted structures. This step is performed in the Dock Library step of the Combinatorial Screening panel, and is described in detail in Chapter 8.

10. Select the optimal library.

    With the docking stage complete, you can focus the library using the docking results to select the reagents that yield the best-scoring structures, and filter the results based on a selection of ADME properties. Finally, you can enumerate, and optionally dock, the focused combinatorial library. These tasks are performed in the Analyze Library step of the Combinatorial Screening panel. This step is described in detail in Chapter 9.

1.3 Overview of Library Enumeration

If you already have a focused set of reagents, or if you simply want to enumerate a library, you can run a combinatorial library enumeration job. You must obtain and prepare the reagents in the same way as for library design (Step 2 and Step 5); and you must prepare the core, define the attachment positions and associate the reagent files with the attachment positions (Step 3 and Step 6). The attachments are defined and the library enumerated in the Combinatorial Library Enumeration panel, which is described in Chapter 10.

1.4 Citing CombiGlide in Publications

The use of this program should be acknowledged in publications as:

CombiGlide, version 1.0, Schrödinger, LLC, New York, NY, 2005.

Introduction to Maestro

Maestro is the graphical user interface for all of Schrödinger's products: CombiGlide™, Epik™, Glide™, Impact™, Jaguar™, Liaison™, LigPrep™, MacroModel®, Phase™, Prime™, QikProp™, QSite™, SiteMap™, and Strike™. It contains tools for building, displaying, and manipulating chemical structures; for organizing, loading, and storing these structures and associated data; and for setting up, monitoring, and visualizing the results of calculations on these structures. This chapter provides a brief introduction to Maestro and some of its capabilities. For more information on any of the topics in this chapter, see the *Maestro User Manual*.

2.1 General Interface Behavior

Most Maestro panels are amodal: more than one panel can be open at a time, and a panel need not be closed for an action to be carried out. Each Maestro panel has a Close button so you can hide the panel from view.

Maestro supports the mouse functions common to many graphical user interfaces. The left button is used for choosing menu items, clicking buttons, and selecting objects by clicking or dragging. This button is also used for resizing and moving panels. The right button displays a shortcut menu. Other common mouse functions are supported, such as using the mouse in combination with the SHIFT or CTRL keys to select a range of items and select or deselect a single item without affecting other items.

In addition, the mouse buttons are used for special functions described later in this chapter. These functions assume that you have a three-button mouse. If you have a two-button mouse, ensure that it is configured for three-button mouse simulation (the middle mouse button is simulated by pressing or holding down both buttons simultaneously).

2.2 Starting Maestro

Before starting Maestro, you must first set the `SCHRODINGER` environment variable to point to the installation directory. To set this variable, enter the following command at a shell prompt:

csh/tcsh: `setenv SCHRODINGER` *installation-directory*

bash/ksh: `export SCHRODINGER=`*installation-directory*

You might also need to set the DISPLAY environment variable, if it is not set automatically when you log in. To determine if you need to set this variable, enter the command:

```
echo $DISPLAY
```

If the response is a blank line, set the variable by entering the following command:

csh/tcsh: `setenv DISPLAY` *display-machine-name*`:0.0`

bash/ksh: `export DISPLAY=`*display-machine-name*`:0.0`

After you set the SCHRODINGER and DISPLAY environment variables, you can start Maestro using the command:

```
$SCHRODINGER/maestro options
```

If you add the $SCHRODINGER directory to your path, you only need to enter the command maestro. Options for this command are given in Section 2.1 of the *Maestro User Manual*.

The directory from which you started Maestro is Maestro's current working directory, and all data files are written to and read from this directory unless otherwise specified (see Section 2.8 on page 27). You can change directories by entering the following command in the command input area (see page 8) of the main window:

```
cd directory-name
```

where *directory-name* is either a full path or a relative path.

2.3 The Maestro Main Window

The Maestro main window is shown in Figure 2.1 on page 7. The main window components are listed below.

The following components are always visible:

- **Title bar**—displays the Maestro version, the project name (if there is one) and the current working directory.
- **Auto-Help**—automatically displays context-sensitive help.
- **Menu bar**—provides access to panels.
- **Workspace**—displays molecular structures and other 3D graphical objects.

The following components can be displayed or hidden by choosing the component from the Display menu. Your choice of which main window components are displayed is persistent between Maestro sessions.

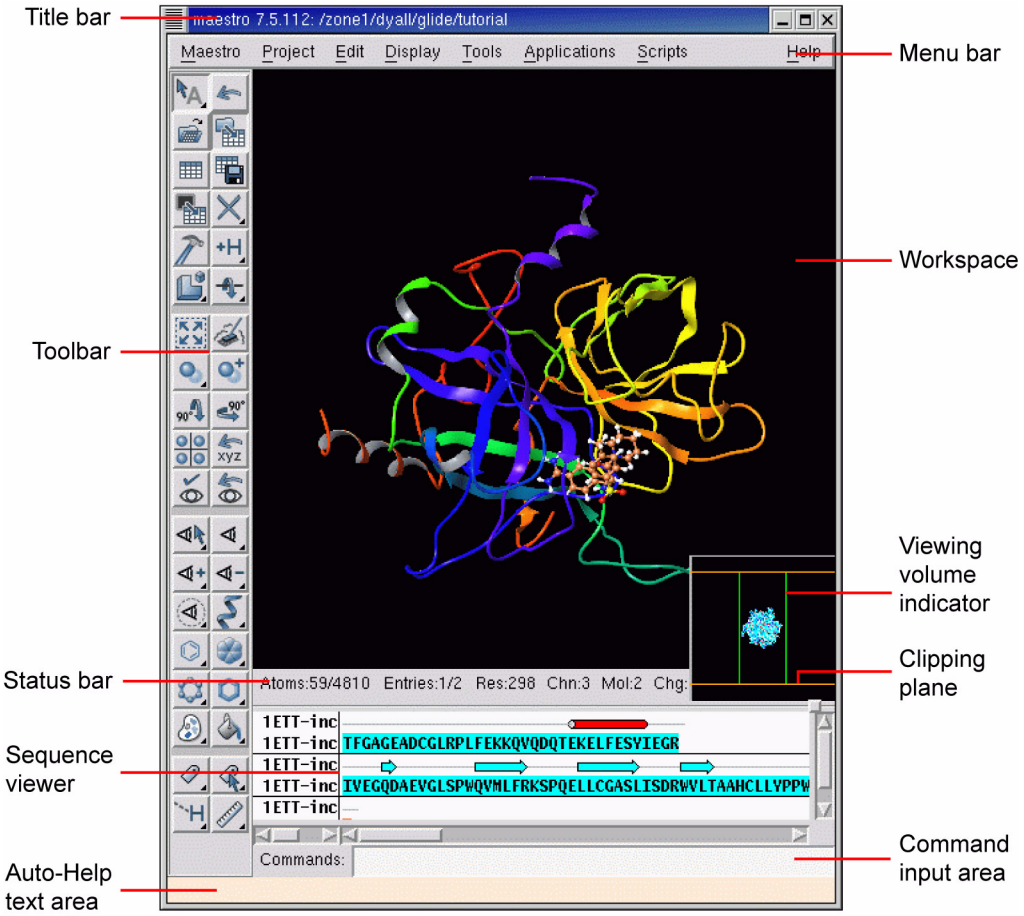


Figure 2.1. The Maestro main window.

- **Toolbar**—contains buttons for many common tasks and provides tools for displaying and manipulating structures, as well as organizing the Workspace.
- **Status bar**—displays information about a particular atom, or about structures in the Workspace, depending on where the pointer pauses (see Section 2.5 of the *Maestro User Manual* for details):
  - **Atom**—displays the chain, residue number, element, PDB atom name, formal charge, and title or entry name (this last field is set by choosing Preferences from the Maestro menu and selecting the Feedback folder).
  - **Workspace**—displays the number of atoms, entries, residues, chains, and molecules in the Workspace.

- **Clipping planes window**—displays a small, top view of the Workspace and shows the clipping planes and viewing volume indicators.
- **Sequence viewer**—shows the sequences for proteins displayed in the Workspace. See Section 2.6 of the *Maestro User Manual* for details.
- **Command input area**—provides a place to enter Maestro commands.

When a distinction between components in the main window and those in other panels is needed, the term *main* is applied to the main window components (e.g., main toolbar).

You can expand the Workspace to occupy the full screen, by pressing CTRL+=. All other components and panels are hidden. To return to the previous display, press CTRL+= again.

2.3.1 The Menu Bar

The menus on the main menu bar provide access to panels, allow you to execute commands, and control the appearance of the Workspace. The main menus are as follows:

- Maestro—save or print images in the Workspace, execute system commands, save or load a panel layout, set preferences, set up Maestro command aliases, and quit Maestro.
- Project—open and close projects, import and export structures, make a snapshot, and annotate a project. These actions can also be performed from the Project Table panel. For more information, see Section 2.4 on page 13.
- Edit—undo actions, build and modify structures, define command scripts and macros, and find atoms in the Workspace.
- Display—control the display of the contents of the Workspace, arrange panels, and display or hide main window components.
- Tools—group atoms; measure, align, and superimpose structures; and view and visualize data.
- Applications—set up, submit, and monitor jobs for Schrödinger's computational programs. Some products have a submenu from which you can choose the task to be performed.
- Scripts—manage and install Python scripts that come with the distribution and scripts that you create yourself. (See Chapter 13 of the *Maestro User Manual* for details.)
- Help—open the Help panel, the PDF documentation index, or information panels; run a demonstration; and display or hide Balloon Help (tooltips).

2.3.2 The Toolbar

The main toolbar contains three kinds of buttons for performing common tasks:

Action—Perform a simple task, like clearing the Workspace.

Display—Open or close a panel or open a dialog box, such as the Project Table panel.

Menu—Display a *button menu*. These buttons have a triangle in the lower right corner.

There are four types of items on button menus, and all four types can be on the same menu (see Figure 2.2):

- **Action**—Perform an action immediately.
- **Display**—Open a panel or dialog box.
- **Object types for selection**—Choose Atoms, Bonds, Residues, Chains, Molecules, or Entries, then click on an atom in the Workspace to perform the action on all the atoms in that structural unit.

  The object type is marked on the menu with a red diamond and the button is indented to indicate the action to be performed.
- **Other setting**—Set a state, choose an attribute, or choose a parameter and click on atoms in the Workspace to display or change that parameter.

The toolbar buttons are described below. Some descriptions refer to features not described in this chapter. See the *Maestro User Manual* for a fuller description of these features.

Figure 2.2. The Workspace selection ***button menu and the*** Adjust distances, angles or dihedrals ***button menu.***

Workspace selection
– Choose an object type for selecting
– Open the Atom Selection dialog box

Undo/Redo
Undo or redo the last action. Performs the same function as the Undo item on the Edit menu, and changes to an arrow pointing in the opposite direction when an Undo has been performed, indicating that its next action is Redo.

Open a project
Open the Open Project dialog box.

Import structures
Open the Import panel.

Open/Close Project Table
Open the Project Table panel or close it if it is open.

Save as
Open the Save Project As dialog box, to save the project with a new name.

Create entry from Workspace
Open a dialog box in which you can create an entry in the current project using the contents of the Workspace.

Delete
– Choose an object type for deletion
– Delete hydrogens and waters
– Open the Atom Selection dialog box
– Delete other items associated with the structures in the Workspace
– Click to select atoms to delete
– Double-click to delete all atoms

Open/Close Build panel
Open the Build panel or close it if it is open.

Add hydrogens
– Choose an object type for applying a hydrogen treatment
– Open the Atom Selection dialog box
– Click to select atoms to treat
– Double-click to apply to all atoms

Local transformation
– Choose an object type for transforming
– Click to select atoms to transform
– Open the Advanced Transformations panel

Adjust distances, angles or dihedrals
– Choose a parameter for adjusting
– Delete adjustments

Fit to screen
Scale the displayed structure to fit into the Workspace and reset the center of rotation.

Clear Workspace
Clear all atoms from the Workspace.

Set fog display state
Choose a fog state. Automatic means fog is on when there are more than 40 atoms in the Workspace, otherwise it is off.

Enhance depth cues
Optimize fogging and other depth cues based on what is in the Workspace.

Rotate around X axis by 90 degrees
Rotate the Workspace contents around the X axis by 90 degrees.

Rotate around Y axis by 90 degrees
Rotate the Workspace contents around the Y axis by 90 degrees.

Tile entries
Arrange entries in a rectangular grid in the Workspace.

Reset Workspace
Reset the rotation, translation, and zoom of the Workspace to the default state.

Save view
Save the current view of the Workspace: orientation, location, and zoom.

Restore view
Restore the last saved view of the Workspace: orientation, location, and zoom.

Display only selected atoms
– Choose an object type for displaying
– Click to select atoms to display
– Double-click to display all atoms

Display only
– Choose a predefined atom category
– Open the Atom Selection dialog box

Also display
– Choose a predefined atom category
– Open the Atom Selection dialog box

Undisplay
– Choose a predefined atom category
– Open the Atom Selection dialog box

Display residues within N angstroms of currently displayed atoms
– Choose a radius
– Open a dialog box to set a value

Show, hide, or color ribbons
– Choose to show or hide ribbons
– Choose a color scheme for coloring ribbons

Draw bonds in wire
– Choose an object type for drawing bonds in wire representation
– Open the Atom Selection dialog box
– Click to select atoms for representation
– Double-click to apply to all atoms

Draw atoms in CPK
– Choose an object type for drawing bonds in CPK representation
– Open the Atom Selection dialog box
– Click to select atoms for representation
– Double-click to apply to all atoms

Draw atoms in Ball & Stick
– Choose an object type for drawing bonds in Ball & Stick representation
– Open the Atom Selection dialog box
– Click to select atoms for representation
– Double-click to apply to all atoms

Draw bonds in tube
– Choose an object type for drawing bonds in tube representation
– Open the Atom Selection dialog box
– Click to select atoms for representation
– Double-click to apply to all atoms

Color all atoms by scheme
Choose a predefined color scheme.

Color residue by constant color
– Choose a color for applying to residues
– Click to select residues to color
– Double-click to color all atoms

Label atoms
– Choose a predefined label type
– Delete labels

Label picked atoms
– Choose an object type for labeling atoms
– Open the Atom Selection dialog box
– Open the Atom Labels panel at the Composition folder
– Delete labels
– Click to select atoms to label
– Double-click to label all atoms

Display H-bonds
– Choose bond type:
intra—displays H-bonds within the selected molecule
inter—displays H-bonds between the selected molecule and all other atoms.
– Delete H-bonds
– Click to select molecule

Measure distances, angles or dihedrals
– Choose a parameter for displaying measurements
– Delete measurements
– Click to select atoms for measurement

2.3.3 Mouse Functions in the Workspace

The left mouse button is used for selecting objects. You can either click on a single atom or bond, or you can drag to select multiple objects. The right mouse button opens shortcut menus, which are described in Section 2.7 of the *Maestro User Manual*.

The middle and right mouse buttons can be used on their own and in combination with the SHIFT and CTRL keys to perform common operations, such as rotating, translating, centering, adjusting, and zooming.

Table 2.1. Mapping of Workspace operations to mouse actions.

Mouse Button	Keyboard	Motion	Action
Left		click, drag	Select
Left	SHIFT	click, drag	Toggle the selection
Middle		drag	Rotate about X and Y axes Adjust bond, angle, or dihedral
Middle	SHIFT	drag vertically	Rotate about X axis
Middle	SHIFT	drag horizontally	Rotate about Y axis
Middle	CTRL	drag horizontally	Rotate about Z axis
Middle	SHIFT + CTRL	drag horizontally	Zoom
Right		click	Spot-center on selection
Right		click and hold	Display shortcut menu
Right		drag	Translate in the X-Y plane
Right	SHIFT	drag vertically	Translate along the X axis
Right	SHIFT	drag horizontally	Translate along the Y axis
Right	CTRL	drag horizontally	Translate along the Z axis
Middle & Right		drag horizontally	Zoom

2.3.4 Shortcut Key Combinations

Some frequently used operations have been assigned shortcut key combinations. The shortcuts available in the main window are described in Table 2.2.

Table 2.2. Shortcut keys in the Maestro main window.

Keys	Action	Equivalent Menu Choices
CTRL+B	Open Build panel	Edit > Build
CTRL+C	Create entry	Project > Create Entry From Workspace
CTRL+E	Open Command Script Editor panel	Edit > Command Script Editor
CTRL+F	Open Find Atoms panel	Edit > Find
CTRL+H	Open Help panel	Help > Help
CTRL+I	Open Import panel	Project > Import Structures
CTRL+M	Open Measurements panel	Tools > Measurements
CTRL+N	Create new project	Project > New
CTRL+O	Open project	Project > Open
CTRL+P	Print	Maestro > Print
CTRL+Q	Quit	Maestro > Quit
CTRL+S	Open Sets panel	Tools > Sets
CTRL+T	Open Project Table panel	Project > Show Table
CTRL+W	Close project	Project > Close
CTRL+Z	Undo/Redo last command	Edit > Undo/Redo
CTRL+=	Enter and exit full screen mode (Workspace occupies full screen)	None

2.4 Maestro Projects

All the work you do in Maestro is done within a *project*. A project consists of a set of *entries*, each of which contains one or more chemical structures and their associated data. In any Maestro session, there can be only one Maestro project open. If you do not specify a project when you start Maestro, a *scratch* project is created. You can work in a scratch project without saving it, but you must save it in order to use it in future sessions. When you save or close a project, all the view transformations (rotation, translation, and zoom) are saved with it. When you close a project, a new scratch project is automatically created.

Likewise, if there is no entry displayed in the Workspace, Maestro creates a *scratch* entry. Structures that you build in the Workspace constitute a scratch entry until you save the structures as project entries. The scratch entry is not saved with the project unless you explicitly add it to the project. However, you can use a scratch entry as input for some calculations.

To add a scratch entry to a project, do one of the following:

- Click the Create entry from Workspace button:

- Choose Create Entry from Workspace from the Project menu.
- Press CTRL+C.

In the dialog box, enter a name and a title for the entry. The entry name is used internally to identify the entry and can be modified by Maestro. The title can be set or changed by the user, but is not otherwise modified by Maestro.

Once an entry has been incorporated into the project, its structures and their data are represented by a row in the Project Table. Each row contains the row number, an icon indicating whether the entry is displayed in the Workspace (the In column), the entry title, a button to open the Surfaces panel if the entry has surfaces, the entry name, and any entry properties. The row number is not a property of the entry.

Entries can be collected into groups, and the members of the group can be displayed or hidden. Most additions of multiple entries to the Project Table are done as entry groups.

You can use entries as input for all of the computational programs—Glide, Impact, Jaguar, Liaison, LigPrep, MacroModel, Phase, Prime, QikProp, QSite, and Strike. You can select entries as input for the ePlayer, which displays the selected structures in sequence. You can also duplicate, combine, rename, and sort entries; create properties; import structures as entries; and export structures and properties from entries in various formats.

To open the Project Table panel, do one of the following:

- Click the Open/Close Project Table button on the toolbar

- Choose Show Table from the Project menu
- Press CTRL+T.

The Project Table panel contains a menu bar, a toolbar, and the table itself.

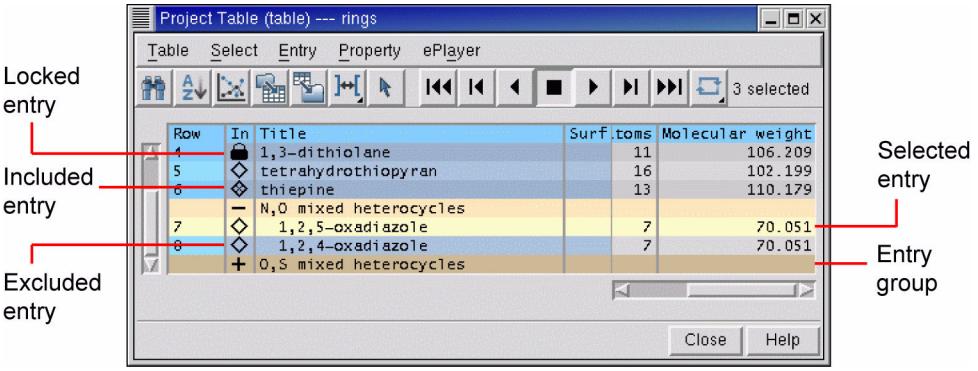


Figure 2.3. The** Project Table **panel.

2.4.1 The Project Table Toolbar

The Project Table toolbar contains two groups of buttons and a status display. The first set of buttons opens various panels that allow you to perform functions on the entries in the Project Table. The second set of buttons controls the ePlayer, which "plays through" the selected structures: each structure is displayed in the Workspace in sequence, at a given time interval. See Section 2.3.2 on page 9 for a description of the types of toolbar buttons. The buttons are described below.

Find
Open the Find panel for locating alphanumeric text in any column of the Project Table, except for the row number.

Sort
Open the Sort panel for sorting entries by up to three properties.

Plot
Open the Plot panel for plotting entry properties.

Import Structure
Open the Import panel for importing structures into the project.

Export Structure
Open the Export panel for exporting structures to a file.

Columns
Choose an option for adjusting the column widths.

Select only
Open the Entry Selection dialog box for selecting entries based on criteria for entry properties.

Go to start
Display the first selected structure.

Previous
Display the previous structure in the list of selected structures.

Play backward
Display the selected structures in sequence, moving toward the first.

Stop
Stop the ePlayer.

Play forward
Display the selected structures in sequence, moving toward the last.

Next
Display the next structure in the list of selected structures.

Go to end
Display the last selected structure.

Loop
Choose an option for repeating the display of the structures. Single Direction displays structures in a single direction, then repeats. Oscillate reverses direction each time the beginning or end of the list is reached.

The status display, to the right of the toolbar buttons, shows the number of selected entries. When you pause the cursor over the status display, the Balloon Help shows the total number of entries, the number shown in the table, the number selected, and the number included in the Workspace.

2.4.2 The Project Table Menus

- Table—find text, sort entries, plot properties, import and export structures, and configure the Project Table.
- Select—select all entries, none, invert your selection, or select classes of entries using the Entry Selection dialog box and the Filter panel.
- Entry—include or exclude entries from the Workspace, display or hide entries in the Project Table, and perform various operations on the selected entries.
- Property—display and manipulate entry properties in the Project Table.
- ePlayer—view entries in succession, stop, reverse, and set the ePlayer options.

2.4.3 Selecting Entries

Many operations in Maestro are performed on the entries selected in the Project Table. The Project Table functions much like any other table: select rows by clicking, shift-clicking, and control-clicking. However, because clicking in an editable cell of a selected row enters edit mode, you should click in the Row column to select entries. See Section 2.4.5 on page 18 for more information on mouse actions in the Project Table. There are shortcuts for selecting classes of entries on the Select menu.

In addition to selecting entries manually, you can select entries that meet a combination of conditions on their properties. Such combinations of conditions are called *filters*. Filters are Entry Selection Language (ESL) expressions and are evaluated at the time they are applied. For example, if you want to set up a Glide job that uses ligands with a low molecular weight (say, less than 300) and that has certain QikProp properties, you can set up a filter and use it to select entries for the job. If you save the filter, you can use it again on a different set of ligands that meet the same selection criteria.

To create a filter:

1. Do one of the following:
   - Choose Only, Add, or Deselect from the Select menu.
   - Click the Entry selection button on the toolbar.

   
2. In the Properties folder, select a property from the property list, then select a condition.
3. Combine this selection with the current filter by clicking Add, Subtract, or Intersect. These buttons perform the Boolean operations OR, AND NOT, and AND on the corresponding ESL expressions.
4. To save the filter for future use click Create Filter, enter a name, and click OK.
5. Click OK to apply the filter immediately.

2.4.4 Including Entries in the Workspace

In addition to selecting entries, you can also use the Project Table to control which entries are displayed in the Workspace. An entry that is displayed in the Workspace is *included* in the Workspace; likewise, an entry that is not displayed is *excluded*. Included entries are marked by an X in the diamond in the In column; excluded entries are marked by an empty diamond. Entry inclusion is completely independent of entry selection.

To include or exclude entries, click, shift-click, or control-click in the In column of the entries, or select entries and choose Include or Exclude from the Entry menu. Inclusion with the mouse works just like selection: when you include an entry by clicking, all other entries are excluded.

It is sometimes useful to keep one entry in the Workspace and include others one by one: for example, a receptor and a set of ligands. You can fix the receptor in the Workspace by selecting it in the Project Table and choosing Fix from the Entry menu or by pressing CTRL+F. A padlock icon replaces the diamond in the In column to denote a *fixed* entry. To remove a fixed entry from the Workspace, you must exclude it explicitly (CTRL+X). It is not affected by the inclusion or exclusion of other entries. Fixing an entry affects only its inclusion; you can still rotate, translate, or modify the structure.

2.4.5 Mouse Functions in the Project Table

The Project Table supports the standard use of shift-click and control-click to select objects. This behavior applies to the selection of entries and the inclusion of entries in the Workspace. You can also drag to resize rows and columns and to move rows.

You can drag a set of non-contiguous entries to reposition them in the Project Table. When you release the mouse button, the entries are placed after the first unselected entry that precedes the entry on which the cursor is resting. For example, if you select entries 2, 4, and 6, and release the mouse button on entry 3, these three entries are placed after entry 1, because entry 1 is the first unselected entry that precedes entry 3. To move entries to the top of the table, drag them above the top of the table; to move entries to the end of the table, drag them below the end of the table.

A summary of mouse functions in the Project Table is provided in Table 2.3.

Table 2.3. Mouse operations in the Project Table.

Task	Mouse Operation
Change a Boolean property value	Click repeatedly in a cell to cycle through the possible values (On, Off, Clear)
Display the Entry menu for an entry	Right-click anywhere in the entry. If the entry is not selected, it becomes the selected entry. If the entry is selected, the action is applied to all selected entries.
Display a version of the Property menu for a property	Right-click in the column header
Edit the text or the value in a table cell	Click in the cell and edit the text or value
Include an entry in the Workspace, exclude all others	Click the In column of the entry

Table 2.3. Mouse operations in the Project Table. (Continued)

Task	Mouse Operation
Move selected entries	Drag the entries
Paste text into a table cell	Middle-click
Resize rows or columns	Drag the boundary with the middle mouse button
Select an entry, deselect all others	For an unselected entry, click anywhere in the row except the In column; for a selected entry, click the row number.
Select or include multiple entries	Click the first entry then shift-click the last entry
Toggle the selection or inclusion state	Control-click the entry or the In column

2.4.6 Project Table Shortcut Keys

Some frequently used project operations have been assigned shortcut key combinations. The shortcuts, their functions, and their menu equivalents are listed in Table 2.4.

Table 2.4. Shortcut keys in the Project Table.

Keys	Action	Equivalent Menu Choices
CTRL+A	Select all entries	Select > All
CTRL+F	Fix entry in Workspace	Entry > Fix
CTRL+I	Open Import panel	Table > Import Structures
CTRL+N	Include only selected entries	Entry > Include Only
CTRL+U	Deselect all entries	Select > None
CTRL+X	Exclude selected entries	Entry > Exclude
CTRL+Z	Undo/Redo last command	Edit > Undo/Redo in main window

2.5 Building a Structure

After you start Maestro, the first task is usually to create or import a structure. You can open existing Maestro projects or import structures from other sources to obtain a structure, or you can build your own. To open the Build panel, do one of the following:

- Click the Open/Close Build panel button in the toolbar:

- Choose Build from the Edit menu.
- Press CTRL+B.

The Build panel allows you to create structures by drawing or placing atoms or fragments in the Workspace and connecting them into a larger structure, to adjust atom positions and bond orders, and to change atom properties. This panel contains a toolbar and three folders.

2.5.1 Placing and Connecting Fragments

The Build panel provides several tools for creating structures in the Workspace. You can place and connect fragments, or you can draw a structure freehand.

To place a fragment in the Workspace:

1. Select Place.
2. Choose a fragment library from the Fragments menu.
3. Click a fragment.
4. Click in the Workspace where you want the fragment to be placed.

To connect fragments in the Workspace, do one of the following:

- Place another fragment and connect them using the Connect & Fuse panel, which you open from the Edit menu on the main menu bar or with the Display Connect & Fuse panel on the Build toolbar.

- Replace one or more atoms in the existing fragment with another fragment by selecting a fragment and clicking in the Workspace on the main atom to be replaced.
- Grow another fragment by selecting Grow in the Build panel and clicking the fragment you want to add in the Fragments folder.

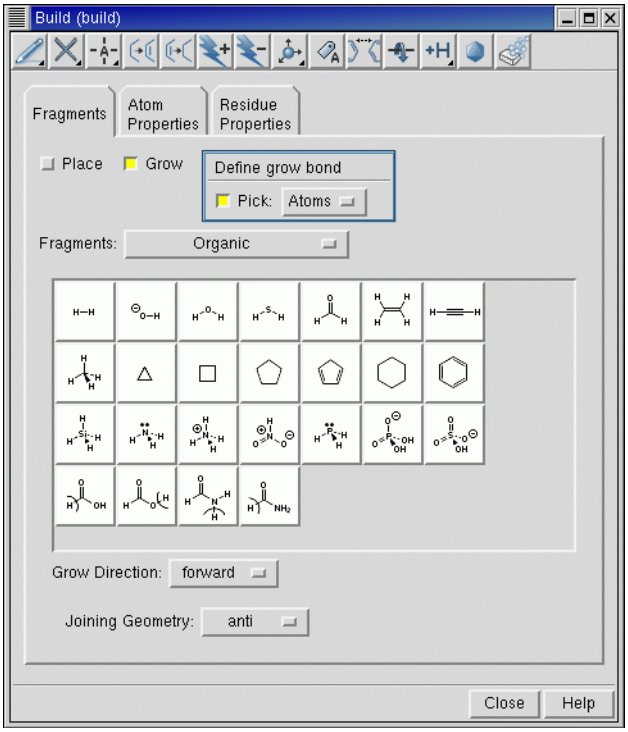


Figure 2.4. The* Build *panel.

Grow mode uses predefined rules to connect a fragment to the *grow bond*. The grow bond is marked by a green arrow. The new fragment replaces the atom at the head of the arrow on the grow bond and all atoms attached to it. To change the grow bond, choose Bonds from the Pick option menu in the Build panel and click on the desired grow bond in the Workspace. The arrow points to the atom nearest to where you clicked.

To draw a structure freehand:

1. Choose an element from the Draw button menu on the Build panel toolbar:

2. Click in the Workspace to place an atom of that element.
3. Click again to place another atom and connect it to the previous atom.
4. Continue this process until you have drawn the structure.
5. Click the active atom again to finish drawing.

2.5.2 Adjusting Properties

In the Atom Properties folder, you can change the properties of the atoms in the Workspace. For each item on the Property option menu—Element, Atom Type (MacroModel), Partial Charge, PDB Atom Name, Grow Name, and Atom Name—there is a set of tools you can use to change the atom properties. For example, the Element tools consist of a periodic table from which you can choose an element and select an atom to change it to an atom of the selected element.

Similarly, the Residue Properties folder provides tools for changing the properties of residues: the Residue Number, the Residue Name, and the Chain Name.

To adjust bond lengths, bond angles, dihedral angles, and chiralities during or after building a structure, use the Adjust distances, angles or dihedrals button on the main toolbar:

You can also open the Adjust panel from this button menu, from the Display Adjust panel button on the Build panel toolbar (which has the same appearance as the above button) or from the Edit menu in the main window.

2.5.3 The Build Panel Toolbar

The toolbar of the Build panel provides quick access to tools for drawing and modifying structures and labeling atoms. See Section 2.3.2 on page 9 for a description of the types of toolbar buttons. The toolbar buttons and their use are described below.

Free-hand drawing
Choose an element for drawing structures freehand in the Workspace (default C). Each click in the Workspace places an atom and connects it to the previous atom.

Delete
Choose an object for deleting. Same as the Delete button on the main toolbar, see page 10.

Set element
Choose an element for changing atoms in the Workspace (default C). Click an atom to change it to the selected element.

Increment bond order
Select a bond to increase its bond order by one, to a maximum of 3.

Decrement bond order
Select a bond to decrease its bond order by one, to a minimum of 0.

Increment formal charge
Select an atom to increase its formal charge by one.

Decrement formal charge
Select an atom to decrease its formal charge by one.

Move
Choose a direction for moving atoms, then click the atom to be moved. Moves in the XY plane are made by clicking the new location. Moves in the Z direction are made in 0.5 Å increments.

Label
Apply heteroatom labels as you build a structure. The label consists of the element name and formal charge, and is applied to atoms other than C and H.

Display Connect & Fuse panel
Open the Connect & Fuse panel so you can connect structures (create bonds between structures) or fuse structures (replace atoms of one structure with those of another).

Display Adjust panel
Open the Adjust panel so you can change bond lengths, bond angles, dihedral angles, or atom chiralities.

Add hydrogens
Choose an atom type for applying the current hydrogen treatment. Same as the Add hydrogens button on the main toolbar, see page 10.

Geometry Symmetrizer
Open the Geometry Symmetrizer panel for symmetrizing the geometry of the structure in the Workspace.

Geometry Cleanup
Clean up the geometry of the structure in the Workspace.

2.6 Selecting Atoms

Maestro has a powerful set of tools for selecting atoms in a structure: toolbar buttons, picking tools in panels, and the Atom Selection dialog box. These tools allow you to select atoms in two ways:

- Select atoms first and apply an action to them
- Choose an action first and then select atoms for that action

2.6.1 Toolbar Buttons

The small triangle in the lower right corner of a toolbar button indicates that the button contains a menu. Many of these buttons allow you to choose an object type for selecting: choose Atoms, Bonds, Residues, Chains, Molecules, or Entries, then click on an atom in the Workspace to perform the action on all the atoms in that structural unit.

For example, to select atoms with the Workspace selection toolbar button:

1. Choose Residues from the Workspace selection button menu:

   The button changes to:

2. Click on an atom in a residue in the Workspace to select all the atoms in that residue.

2.6.2 Picking Tools

The picking tools are embedded in each panel in which you need to select atoms to apply an operation. The picking tools in a panel can include one or more of the following:

- Pick option menu—Allows you to choose an object type. Depending on the operation to be performed, you can choose Atoms, Bonds, Residues, Chains, Molecules, or Entries, then click on an atom in the Workspace to perform the action on all the atoms in that structural unit.

  The Pick option menu varies from panel to panel, because not all object types are appropriate for a given operation. For example, some panels have only Atoms and Bonds in the Pick option menu.

- All button—Performs the action on all atoms in the Workspace.
- Selection button—Performs the action on any atoms already selected in the Workspace.
- Previous button—Performs the action on the most recent atom selection defined in the Atom Selection dialog box.
- Select button—Opens the Atom Selection dialog box.
- ASL text box—Allows you to type in an ASL expression for selecting atoms.

  ASL stands for Atom Specification Language, and is described in detail in the *Maestro Command Reference Manual*.

- Clear button—Clears the current selection

- Show markers option—Marks the selected atoms in the Workspace.

For example, to label atoms with the Label Atoms panel:

1. Choose Atom Labels from the Display menu.
2. In the Composition folder, select Element and Atom Number.
3. In the picking tools section at the top of the panel, you could do one of the following:
   - Click Selection to apply labels to the atoms already selected in the Workspace (from the previous example).
   - Choose Residues from the Pick option menu and click on an atom in a different residue to label all the atoms in that residue.

2.6.3 The Atom Selection Dialog Box

If you wish to select atoms based on more complex criteria, you can use the Atom Selection dialog box. To open this dialog box, choose Select from a button menu or click the Select button in a panel. See Section 5.3 of the *Maestro User Manual* for detailed instructions on how to use the Atom Selection dialog box.

2.7 Scripting in Maestro

Although you can perform nearly all Maestro-supported operations through menus and panels, you can also perform operations using Maestro commands, or compilations of these commands, called *scripts*. Scripts can be used to automate lengthy procedures or repetitive tasks and can be created in several ways. These are summarized below.

2.7.1 Python Scripts

Python is a full-featured scripting language that has been embedded in Maestro to extend its scripting facilities. The Python capabilities within Maestro include access to Maestro functionality for dealing with chemical structures, projects, and Maestro files.

The two main Python commands used in Maestro are:

- `pythonrun`—executes a Python module. (You can also use the alias `pyrun`.) The syntax is:

  `pythonrun` *module*`.`*function*

- `pythonimport`—rereads a Python file so that the next time you use the `pythonrun` command, it uses the updated version of the module. (You can also use the alias `pyimp`.)

From the Maestro Scripts menu you can install, manage, and run Python scripts. For more information on the Scripts menu, see Section 13.1 of the *Maestro User Manual*.

For more information on using Python with Maestro, see *Scripting with Python*.

2.7.2 Command Scripts

All Maestro commands are logged and displayed in the Command Script Editor panel. This means you can create a command script by performing the operations with the GUI controls, copying the logged commands from the Command History list into the Script text area of the panel, then saving the list of copied commands as a script.

To run an existing command script:

1. Open the Command Script Editor panel from the Edit menu in the main window.
2. Click Open Local and navigate to the directory containing the desired script.
3. Select a script in the Files list and click Open.

   The script is loaded into the Script window of the Command Script Editor panel.
4. Click Run Script.

Command scripts cannot be used for Prime operations.

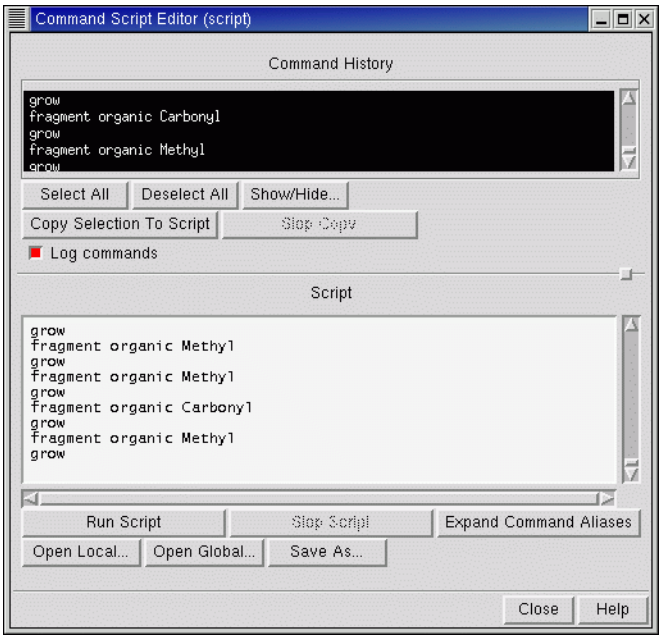


Figure 2.5. The Command Script Editor ***panel.***

2.7.3 Macros

There are two kinds of macros you can create: named macros and macros assigned to function keys F1 through F12.

To create and run a named macro:

1. Open the Macros panel from the Edit menu in the main window.
2. Click New, enter a name for the macro, and click OK.
3. In the Definition text box, type the commands for the macro.
4. Click Update to update the macro definition.
5. To run the macro, enter the following in the command input area in the main window:

   `macrorun` *macro-name*

   If the command input area is not visible, choose Command Input Area from the Display menu.

To create and run a function key macro:

1. Open the Function Key Macros panel from the Edit menu in the main window.
2. From the Macro Key option, select a function key (F1 through F12) to which to assign the macro.
3. In the text box, type the commands for the macro.
4. Click Run to test the macro or click Save to save it.
5. To run the macro from the main window, press the assigned function key.

For more information on macros, see Section 13.5 of the *Maestro User Manual*.

2.8 Specifying a Maestro Working Directory

When you use Maestro to launch CombiGlide jobs, Maestro writes job output to the directory specified in the Directory folder of the Preferences panel. By default, this directory (the file I/O directory) is the directory from which you started Maestro.

To change the Maestro working directory:

1. Open the Preferences panel from the Maestro menu.
2. Click the Directory tab.
3. Select the directory you want to use for reading and writing files.

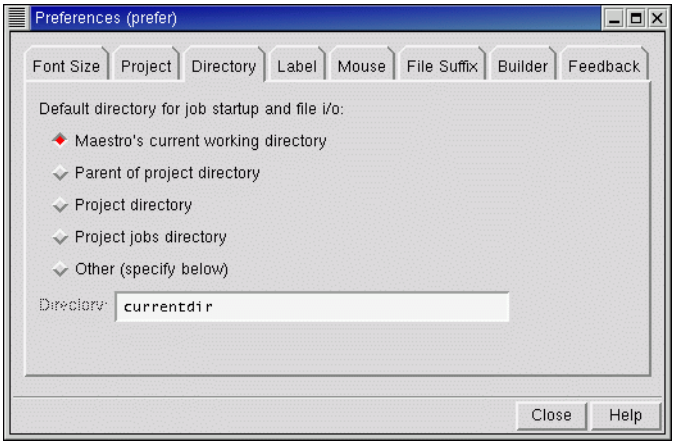


Figure 2.6. The** Directory **folder of the** Preferences **panel.

You can also set other preferences in the Preferences panel. See Section 12.2 of the *Maestro User Manual* for details.

2.9 Undoing an Operation

To undo a single operation, click the Undo button in the toolbar, choose Undo from the Edit menu, or press CTRL+Z. The word Undo in the menu is followed by text that describes the operation to undo. Not all operations can be undone: for example, global rotations and translations are not undoable operations. For such operations you can use the Save view and Restore view buttons in the toolbar, which save and restore a molecular orientation.

2.10 Running and Monitoring Jobs

Maestro has panels for each product for preparing and submitting jobs. To use these panels, choose the appropriate product and task from the Applications menu and its submenus. Set the appropriate options in the panel, then click Start to open the Start dialog box and set options for running the job. For a complete description of the Start dialog box associated with your computational program, see your product's User Manual. When you have finished setting the options, click Start to launch the job and open the Monitor panel.

The Monitor panel is the control panel for monitoring the progress of jobs and for pausing, resuming, or killing jobs. All jobs that belong to you can be displayed in the Monitor panel, whether or not they were started from Maestro. Subjobs are indented under their parent in the job list. The text pane shows output information from the monitored job, such as the contents

of the log file. The Monitor panel opens automatically when you start a job. If it is not open, you can open it by choosing Monitor from the Applications menu in the Maestro main window.

While jobs are running, the Detach, Pause, Resume, Stop, Kill, and Update buttons are active. When there are no jobs currently running, only the Monitor and Delete buttons are active. These buttons act on the selected job. By default, only jobs started from the current project are shown. To show other jobs, deselect Show jobs from current project only.

When a monitored job ends, the results are incorporated into the project according to the settings used to launch the job. If a job that is not currently being monitored ends, you can select it in the Monitor panel and click Monitor to incorporate the results. Monitored jobs are incorporated only if they are part of the current project. You can monitor jobs that are not part of the current project, but their results are not incorporated. To add their results to a project, you must open the project and import the results.

Further information on job control, including configuring your site, monitoring jobs, running jobs, and job incorporation, can be found in the *Job Control Guide* and the *Installation Guide*.

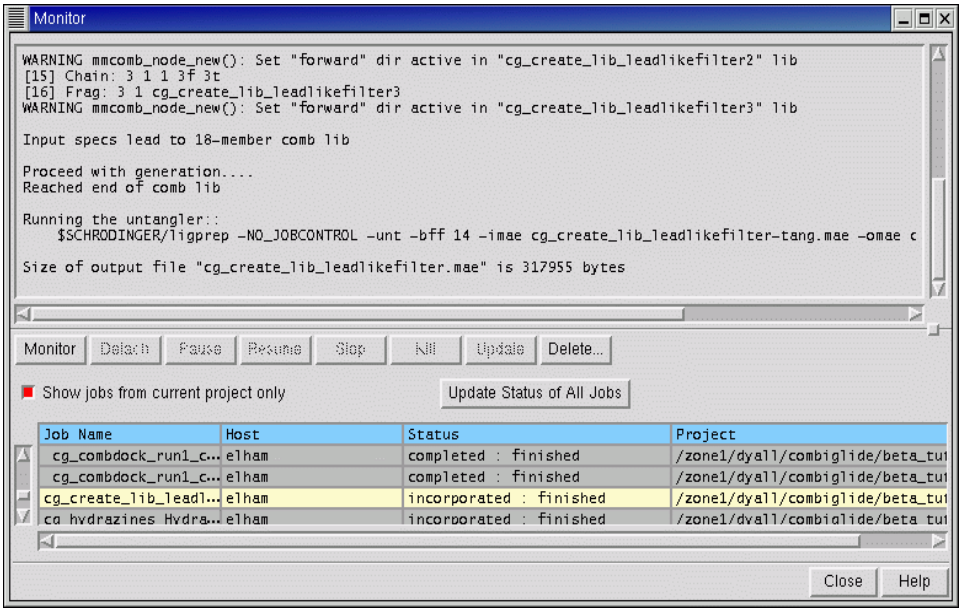


Figure 2.7. The* Monitor *panel.

2.11 Getting Help

Maestro comes with automatic, context-sensitive help (Auto-Help), Balloon Help (tooltips), an online help facility, and a user manual. To get help, follow the steps below:

- Check the Auto-Help text box at the bottom of the main window. If help is available for the task you are performing, it is automatically displayed there. It describes what actions are needed to perform the task.
- If your question concerns a GUI element, such as a button or option, there may be Balloon Help for the item. Pause the cursor over the element. If the Balloon Help does not appear, check that Show Balloon Help is selected in the Help menu of the main window. If there is Balloon Help for the element, it appears within a few seconds.
- If you do not find the help you need using either of the steps above, click the Help button in the lower right corner of the appropriate panel. The Help panel is displayed with a relevant help topic.
- For help with a concept or action not associated with a panel, open the Help panel from the Help menu or press CTRL+H.

If you do not find the information you need in the Maestro help system, check the following sources:

- The *Maestro User Manual*
- The Frequently Asked Questions page on the Schrödinger Support Center.

You can also contact Schrödinger by e-mail or phone for help:

- E-mail: help@schrodinger.com
- Phone: (503) 299-1150

2.12 Ending a Maestro Session

To end a Maestro session, choose Quit from the Maestro menu. To save a log file with a record of all operations performed in the current session, click Quit, save log file in the Quit panel. This information can be useful to Schrödinger support staff when responding to any problem you report.

Defining the Chemistry

The first step in the design of a focused library using CombiGlide is to identify the chemistry associated with desired library. Most likely, this will be done in collaboration with one of your medicinal or combinatorial chemists.

For example, if you were interested in generating a focused pyrazole library in an attempt to identify inhibitors of p38 MAP kinase, your synthetic chemists might suggest the approach described in Figure 3.1.

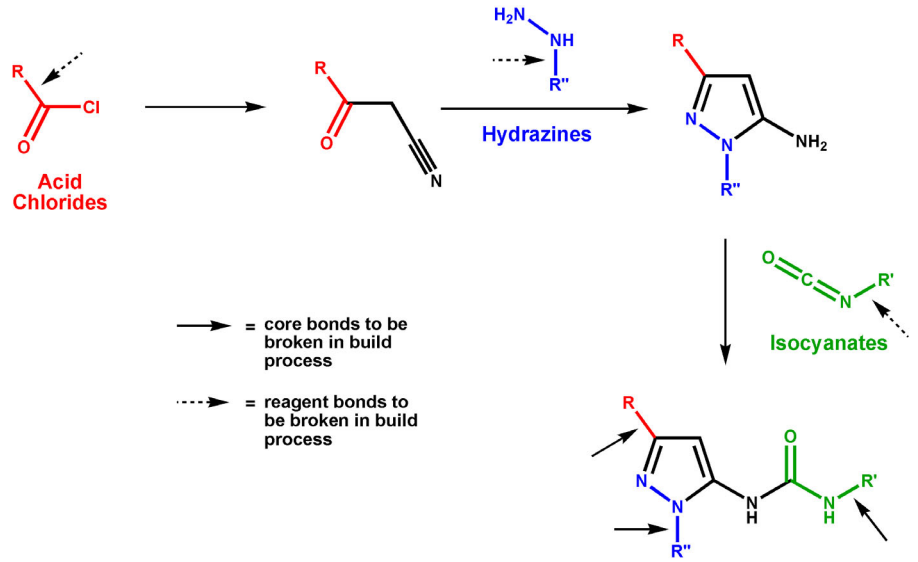


Figure 3.1. Synthetic route to pyrazole library.

3.1 Obtaining Reagent and Protein Structure Files

Once the synthetic approach has been determined, it will be apparent which types of reagents will be required for the library synthesis. For example, in the pyrazole library described in Figure 3.1, the library will be generated by the combination of acid chlorides, hydrazines, and isocyanates.

You must then obtain files that contain the structures of the reagents that you want to use. For each reagent, the structures must be contained in a single, multistructure file. You can provide

files that contain only the structures for particular reagents, or you can provide files that contain structures for a variety of reagents, from which CombiGlide will select out the specified types of reagents. In the pyrazole library example, you could supply one file containing all of the acid chlorides to be considered, one file for the hydrazines, and one for the isocyanates; or you could supply one large file containing structures for all three reagent types. The reagent files must be in either SD or Maestro format.

It is highly recommended that the reagent files contain a set of reagents that has been prefiltered to remove redundant compounds and any reagents that are not compatible with the synthetic approach. Prefiltering will reduce the amount of computational work needed to obtain a focused library and is more likely to lead to useful results.

If you do not already have a library of reagents, you can find links to sources of compounds at the HTScreening.net web site, http://www.htscreening.net/home[1], or the ZINC web site, http://blaster.docking.org/zinc[1]. Most vendors supply structures of the compounds in 2D SD format.

You will also need to obtain a file containing the structure of the receptor. Protein structures can be obtained from the Protein Data Bank (PDB), http://www.rcsb.org/pdb[1]. If you have the PDB database and Prime installed, and the environment variable `SCHRODINGER_PDB` is set, you can import a protein structure from the database directly into Maestro from the Import panel. See Section 3.1 of the *Maestro User Manual* for more information.

3.2 Choosing the Core-Containing Molecule

CombiGlide works by building side chains onto defined positions on a core-containing structure. You must therefore supply a structure that contains the core in addition to supplying the reagents. The reagent structures are used as a source of side chains. Even though in the actual chemical reactions, atoms from the reagents can be included into the core, these atoms are discarded when the side chains are added to the core.

There are some limitations on how you can select the reagents and attachment positions, which may affect the choices you make for the core-containing molecule.

- The bond that is replaced in the core-containing molecule and in the reagent when the side chain is added must be a single bond. In the pyrazole example in Figure 3.1, the bond that is replaced in the isocyanate is the single C–R' bond. The C=N bond in the isocyanate could not be chosen because it is a double bond, even though the corresponding bond in the core-containing molecule is a single bond.
- The bond must not be in a ring. The hydrazine N–N bond in the pyrazole example could not be chosen because it forms part of the pyrazole ring.

1. Please see the notice regarding third party programs and third party Web sites on the copyright page at the front of this manual.

Core-containing molecules are also docked to define core poses, which are used to build and dock the substituted structures. By default, the minimally capped core molecule is used, but you can choose other molecules to dock for the core poses. When you choose core-containing molecules for use in defining core poses, it is often advantageous to select ones that are not so small that they can dock in poses that are unreasonable for the final substituted molecules. It does not matter how long or short the side chains are in the core-containing molecules, since they will be replaced. A good choice would be a known active.

Preparing Structures for CombiGlide

When you have decided on the biological target and the synthetic route to the library you want to generate, and obtained structure files for the reagents and the receptor, you can proceed to the preparation of the structures for CombiGlide. Both the protein structure and the reagent structures must be properly prepared. You can perform these tasks in the Protein Preparation and Reagent Preparation panels. Following protein preparation, you must set up grids for Glide docking, which you can do in the Receptor Grid Generation panel.

4.1 Protein Preparation and Grid Generation

A prepared protein structure is required for CombiGlide docking. Before running the protein preparation job, you must perform some preliminary steps, which may include simplifying any multimeric complexes, deleting unwanted waters or cofactors, correcting serious errors in the protein structure, or adjusting bond orders, ionization states, and formal charges of the protein, metal ions, cofactors, and ligand. The protein preparation job adds hydrogens, neutralizes appropriate amino acid chains, reorients side-chain hydroxyl and thiol groups, and relieves steric clashes. The full procedure for protein preparation is described in Chapter 4 of the *Glide User Manual*. You should refer to this chapter and follow the procedures in it to prepare your protein for CombiGlide. You can open the Protein Preparation panel from the CombiGlide submenu of the Applications menu. This panel is identical to the Glide version of the panel.

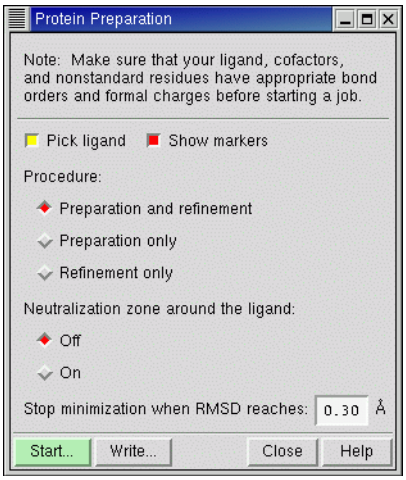


Figure 4.1. The Protein Preparation panel.

When you have prepared the protein, you must generate grids for docking with CombiGlide, using the Receptor Grid Generation panel. The grid generation process is described in detail in Chapter 6 of the *Glide User Manual*. You should refer to this chapter and follow the procedures in it to generate your grids. You can open the Receptor Grid Generation panel from the CombiGlide submenu of the Applications menu. This panel is also identical to the Glide version of the panel.

4.2 Reagent Preparation

Reagent preparation ensures that the input structures are all-atom, 3D structures, and that they have the appropriate information stored with them to construct the molecules that are docked or used for library enumeration. The main tasks in the reagent preparation process are to select the source of reagent structures, select a reagent type (a functional group), identify the bond in the functional group that is replaced when the reagent is added to the core, and perform 2D-to-3D conversion, structure variation and cleanup. The first three of these tasks are performed in the Reagent Preparation panel, and options for the fourth are set up in this panel. A job is then run to obtain the structures from the input file and prepare them for use with CombiGlide.

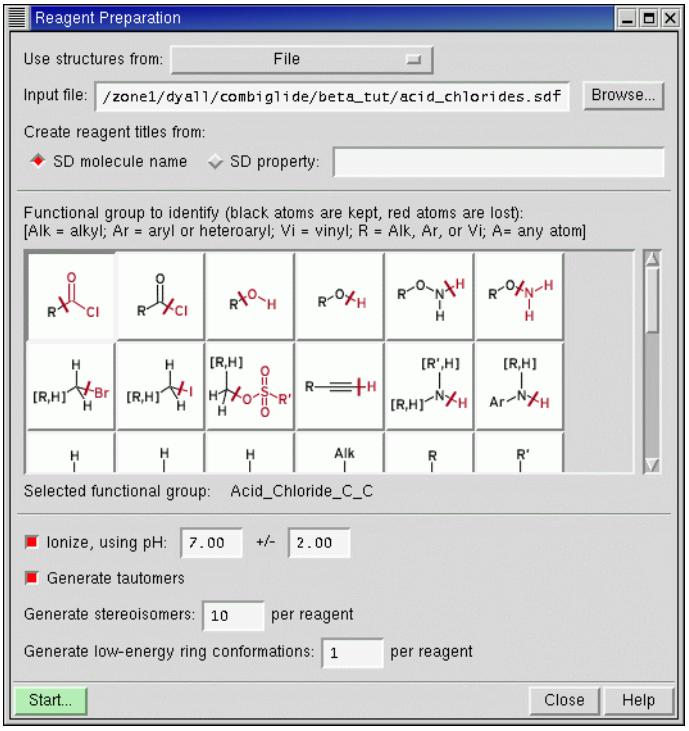


Figure 4.2. The Reagent Preparation panel.

To open the panel, choose Reagent Preparation from the CombiGlide submenu of the Applications menu. This panel contains three sections: one for selecting the input structure file, one for selecting the functional group, and one for setting options for structure conversion, variation, and cleanup.

You must run a reagent preparation job for each reagent type that will be used in combinatorial screening or in combinatorial library enumeration. A reagent type corresponds more or less to a functional group, plus information on which bond is to be replaced in the build process. The tasks described briefly above and in detail below must be performed for each reagent type that you plan to use. In the pyrazole example from Chapter 3, there are three reagent types: acid chlorides, hydrazines, and isocyanates.

4.2.1 Choosing the Source of Reagent Structures

The first task is to choose the source of the reagent structures for a given reagent type. Because the reagent preparation job matches the selected functional group in each structure to a SMARTS pattern, structures that do not match are filtered out. The structures that match are written to a file and passed on to the next stage. The structure input file can therefore contain any structures, not just those that match the desired functional group. However, it is highly recommended that the input file you choose is prefiltered to eliminate redundant structures and structures that are not compatible with the synthetic approach.

There are some restrictions on the structures that you can use as input. If a structure contains more than one instance of the selected functional group, it will not be written to the output file unless it is symmetrical, in which case only one of the bonds will be designated as the bond to be broken in the process of building the library member from the reagents.

Once you have opened the Reagent Preparation panel, you can choose the source of structures from the Use structures from option menu. You can select structures from the Workspace, the Project Table, or from an external file. Structures from the Workspace or the Project Table are written to a Maestro file that is passed to the reagent preparation job.

If you choose File from the Use structures from option menu, you can enter the path to the file in the Input file text box, or click Browse to navigate to the file. The file must be in Maestro format or SD format. The structures can be 2D or 3D, and are converted to all-atom 3D structures by the reagent preparation job.

If you selected an SD file, you can also specify the source of the reagent titles. Reagent titles are used for identification in CombiGlide, so it is important to have unique titles. If you do not make them unique, CombiGlide will do so for you. You can select SD molecule name to use the molecule name from the SD file, or select SD property and enter a property name in the text box to use the property for the reagent titles.

4.2.2 Selecting the Reagent Type

Selecting the reagent type involves choosing a functional group along with the bond that is to be replaced when the reagent is used to build structures for the library. A variety of predefined functional groups with corresponding bonds to be broken are displayed as a set of buttons in the center of the Reagent Preparation panel. The icon for each functional group marks the bond that is broken to attach the reagent to the core with a red line across the bond. The part of the reagent that is discarded is shown in red.

To select a reagent type, click the button that corresponds to the desired combination of a functional group and a bond to break. The selection is displayed below the scrolling region. The "short name" and the "long name" of the selected functional group are displayed below the icon area. Both names consist of a part that identifies the type of compound (acid chloride, primary amine, and so on) and a part that identifies the bond that is broken, separated by an underscore character. The part that identifies the bond consists of two element names separated by an underscore. The first name defines the atom that is kept when the reagent is added to the core, and the second name defines the atom that is discarded, along with anything attached to it. For example, Acid_Chloride_C_C is the long name for acid chlorides in which the carbon-carbon bond of the acid chloride group is broken. In this case the entire –COCl group is discarded, and the R group is added to the core.

A single class of compounds can have multiple icons, corresponding to different bonds that are broken. You should make sure that you have selected the icon with the bond to be broken that is the most appropriate for your molecular system. For example, acid chlorides can have the C–C bond or the C–Cl bond broken.

A detailed description of each functional group is provided in Table 4.1. The functional groups are defined using SMARTS patterns.

If the functional group that you want to use is not listed, you can add a custom functional group. Details on the requirements and procedure are given in Appendix A.

4.2.3 Setting Options for Structure Variation and Cleanup

In addition to defining the atoms of the reagents that will become the side chains, the reagent preparation process performs a 2D-to-3D conversion, generates all reasonable ionization and tautomeric states, and expands the stereochemistry. If the structures you have are already 3D structures in the appropriate ionization state, tautomeric and stereochemical form, you might not need to generate variations. Many collections of reagents are provided in 2D form, with implicit hydrogen atoms. To generate 3D, all-atom structures for docking, some variation to obtain structures that represent the actual state under physiological conditions may be needed.

Table 4.1. Predefined functional groups with definitions

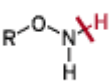
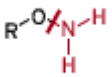
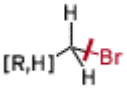
Structure	Short name Long name	Definition of R, R', R", Alk, Ar, Vi, A
	Acid_Cl_C_C Acid_Chloride_C_C	R can be anything with a carbon attached to the carbonyl.
	Acid_Cl_C_Cl Acid_Chloride_C_Cl	See above
	Alc_C_O Alcohol_C_O	R can be an alkyl or aryl group. R cannot have a carbonyl attached to the oxygen of the alcohol.
	Alc_O_H Alcohol_O_H	See above
	Alkoxylamine_N_H Alkoxylamine_N_H	R can be H or anything with a carbon attached to the oxygen of the alkoxylamine.
	Alkoxylamine_O_N Alkoxylamine_O_N	See above
	Alk_Br_C_Br Alkyl_Bromide_C_Br	R can be almost anything: H, alkyl, aryl, alkenyl, alkynyl; alkoxy, aryloxy, alkoxycarbonyl, or aryloxycarbonyl with oxygen attached to CH_2; silyl with silicon attached to CH_2; alkylamino, arylamino, alkylaminocarbonyl, or arylaminocarbonyl with nitrogen attached to CH_2; alkylthio, arylthio, alkylsulfinyl, arylsulfinyl, alkylsufonyl, or arylsulfonyl with sulfur attached to CH_2; ketone with carbonyl attached to CH_2; cyano. R cannot be chloro, iodo, chlorocarbonyl (carbonyl attached to CH_2).

Table 4.1. Predefined functional groups with definitions

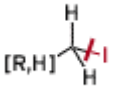
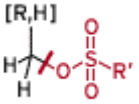
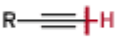
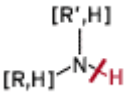
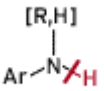
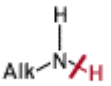
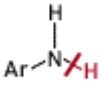
Structure	Short name Long name	Definition of R, R', R", Alk, Ar, Vi, A
	Alk_I_C_I Alkyl_Iodide_C_I	Same as for alkyl bromides
	Alk_Sulf_C_O Alkyl_Sulfonate_C_O	Same as for alkyl bromides
	Any_I_A_I Any_Iodide_A_I	A can be any atom for which force fields are available.
	Alkyne_C_H Alkyne_C_H	R can be H, alkyl, aryl, silyl.
	Amine_Gen_N_H Amine_General_N_H	R can be H, alkyl, aryl. R cannot have a carbonyl carbon attached to the nitrogen of the amine. R' can be H, alkyl, aryl. R' cannot have a carbonyl carbon attached to the nitrogen of the amine.
	Amine_Gen_Ar_N_H Amine_General_Aryl_N_H	Ar can be aryl. R can be H, alkyl, aryl. R cannot have a carbonyl attached to the nitrogen of the amine.
	Amine_Prim_Alk_N_H Amine_Primary_Alkyl_N_H	R can be H, alkyl. R cannot have a carbonyl attached to the nitrogen of the amine.
	Amine_Prim_Ar_N_H Amine_Primary_Aryl_N_H	Ar can be aryl.

Table 4.1. Predefined functional groups with definitions

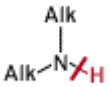
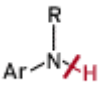
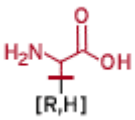
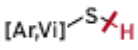
Structure	Short name Long name	Definition of R, R', R", Alk, Ar, Vi, A
	Amine_Prim_Gen_N_H Amine_Primary_General_N_H	R can be alkyl, aryl. R cannot have a carbonyl carbon attached to the nitrogen of the amine.
	Amine_Sec_Alk_N_H Amine_Secondary_Alkyl_N_H	R can be alkyl. R cannot have a carbonyl carbon attached to the nitrogen of the amine. R' can be alkyl. R' cannot have a carbonyl carbon attached to the nitrogen of the amine.
	Amine_Sec_Ar_N_H Amine_Secondary_Aryl_N_H	Ar can be aryl. R can be alkyl, aryl. R cannot have a carbonyl carbon attached to the nitrogen of the amine.
	Amine_Sec_Gen_N_H Amine_Secondary_General_N_H	R can be alkyl, aryl. R cannot have a carbonyl carbon attached to the nitrogen of the amine. R' can be alkyl, aryl. R' cannot have a carbonyl carbon attached to the nitrogen of the amine.
	A_Acid_C_C Amino_Acid_C_C	R can be H, alkyl, aryl.
	Ar_or_Vinyl_Br_C_Br Aryl_or_Vinyl_Bromide_C_Br	Ar is an aryl group; Vi is a vinyl (C=C) group. The aryl or vinyl group must be directly attached to the Br.
	Ar_or_Vinyl_I_C_I Aryl_or_Vinyl_Iodide_C_I	Ar is an aryl group; Vi is a vinyl (C=C) group. The aryl or vinyl group must be directly attached to the I.
	Ar_or_Vinyl_SH_S_H Aryl_or_Vinyl_Thiol_S_H	Ar is an aryl group; Vi is a vinyl (C=C) group. The aryl or vinyl group must be directly attached to the S.

Table 4.1. Predefined functional groups with definitions

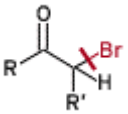
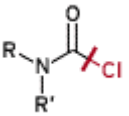
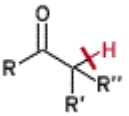
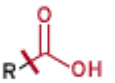
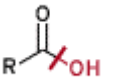
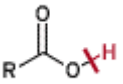
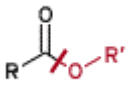
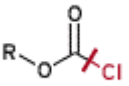
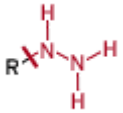
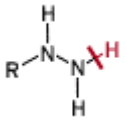
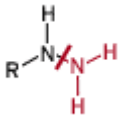
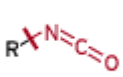
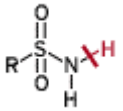
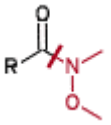
Structure	Short name Long name	Definition of R, R', R", Alk, Ar, Vi, A
	alpha_Br_CO_C_Br alphaBromocarbonyl_C_Br	R can be alkyl, aryl. R' can be H, alkyl, aryl.
	Carbam_Cl_C_Cl Carbamoyl_Chloride_C_Cl	R can be anything with a carbon attached to nitrogen. R' can be anything with a carbon attached to nitrogen.
	alphaCarbonyl_C_H alphaCarbonyl_C_H	R can be H, alkyl, aryl, alkoxy with oxygen attached to carbonyl. R' can be H, alkyl, aryl, carbonyl with carbonyl carbon attached to CH, cyano. R" can be H, alkyl, aryl, carbonyl with carbonyl carbon attached to CH, cyano.
	C_Acid_C_C Carboxylic_Acid_C_C	R can be anything with carbon attached to carbonyl.
	C_Acid_C_O Carboxylic_Acid_C_O	See above
	C_Acid_O_H Carboxylic_Acid_O_H	See above
	C_Ester_C_O Carboxylic_Acid_Ester_C_O	R can be anything with carbon attached to carbonyl. R' can be anything with carbon attached to the oxygen except for a carbonyl carbon.
	Cl_Formate_C_Cl Chloroformate_C_Cl	R can be anything with carbon attached to oxygen.

Table 4.1. Predefined functional groups with definitions

Structure	Short name Long name	Definition of R, R', R", Alk, Ar, Vi, A
	Hydrazine_C_N Hydrazine_C_N	R can be H or anything with a carbon attached to the nitrogen of the hydrazine.
	Hydrazine_N_H Hydrazine_N_H	See above
	Hydrazine_N_N Hydrazine_N_N	See above
	Isocyanate_C_N Isocyanate_C_N	R can be anything with a carbon attached to the nitrogen of the NCO.
	Sulfonamide_N_H Sulfonamide_N_H	R can be anything with a carbon attached to the sulfur of the SO_2.
	Sulf_Cl_S_Cl Sulfonyl_Chloride_S_Cl	R can be anything with a carbon attached to the sulfur of the SO_2.
	Thiol_S_H Thiol_S_H	R can be H, alkyl, aryl, vinyl. R cannot have the carbon of a carbonyl attached to the sulfur.
	W_amide_C_N Weinreb_Amide_C_N	R can be alkyl, aryl. R cannot have a carbonyl carbon attached to the amide carbonyl.

The lower portion of the Reagent Preparation panel allows you some control over the structure variation and cleanup process, by setting options for the following:

- Ionization state—You can request that the structure be protonated or deprotonated to represent the actual forms in water in the given pH range, by selecting Ionize, using pH and entering the pH value and range in the accompanying text boxes. If this option is not selected, the ionization state is left as is.
- Tautomerization state—You can request the generation of tautomers of the input structures by selecting Generate tautomers. Ionized structures might have a different tautomeric form from the un-ionized structures.
- Stereochemistry—Stereoisomers of the input structures are generated in the absence of chirality information, either explicitly specified or deducible from the 3D structure. If the input structures are 2D and have chiral centers but do not have chirality information, you should ensure that the number specified in the Generate stereoisomers text box covers the possible combinations.
- Ring conformations—If the input structures contain rings that can exist in more than one low-energy conformation, enter the number of likely low-energy conformations in the Generate low-energy ring conformations text box. Ring conformations are not searched during docking, so the conformations must be set up during the reagent preparation process.

If you do not need to generate structural variations, you should deselect Ionize, using pH and Generate tautomers, and enter `1` in the Generate stereoisomers and Generate low-energy ring conformations text boxes.

4.2.4 Running the Reagent Preparation Job

When you have completed the tasks above, click Start. The Start dialog box opens, in which you can select job options and start the job. When you click Start in the Start dialog box, the Monitor panel opens, and displays information on the progress of the job.

The output structure file contains the prepared reagents. There might be several output structures per input structure if the molecules can exist in several low-energy forms. Consequently, there can be multiple structures for a given combination of reagents on a core (a “compound”) and the final number of structures in the library can be greater than the final number of compounds.

The following output files are required for later use:

jobname.bld	Reagent structures in all modifications labeled for building
jobname.log	Overall log file for the job
jobname.sqlite	Sqlite database built from the .bld file
*jobname*_vpost-stats.txt	Text file giving summary statistics of molecules in the .bld file

4.2.5 Task Summary

To set up and run a reagent preparation job:

1. Choose the source of reagent structures
2. Select the functional group
3. Select 2D-to-3D conversion and structure variation (LigPrep) options
4. Click Start

4.3 Preparing the Core-Containing Molecules

The structures that you select for the core-containing molecules must be all-atom, minimized, 3D structures. If the structure you have does not meet these criteria, you can prepare it using LigPrep. For more information, see the *LigPrep User Manual*.

Note that only the ionization state and tautomer that you select is used for the core-containing molecule when you define the core in the Define Combinations step of the Combinatorial Screening panel or in the Combinatorial Library Enumeration panel. If you want libraries with multiple states of the core, you must generate the states of the core-containing molecule and select them one-at-a-time to generate a library based on each state.

Combinatorial Screening

The goal of combinatorial screening is to create a small combinatorial library of structures that are likely to have high activity towards to a selected target. The library is screened, first by docking to the target receptor and eliminating structures that do not dock well, then by analyzing the docked structures to define a small reagent set that is likely to have a large number of actives among the compounds that are generated.

Once you have decided on the chemistry, obtained files containing the reagents, the receptor and the core-containing molecule, and prepared the structures in these files, you can proceed with the combinatorial screening process. This process is managed from the Combinatorial Screening panel. To open the Combinatorial Screening panel, choose Combinatorial Screening from the CombiGlide submenu of the Applications menu in the main window.

5.1 The Combinatorial Screening Panel

The Combinatorial Screening panel is designed like a wizard, with five steps. To perform a combinatorial screening, you proceed through each step in turn. You do not have to complete all steps in a single session. The design allows you to exit at an intermediate stage and pick up the process again later at any step for which you have the required data.

Each set of data defined by proceeding through the steps in this panel is called a *run*. The results for each run are stored as a separate entity. You can open existing runs, and create and save runs from the File menu. Runs are saved in the Maestro project, so if you want to keep your results, you should create a named Maestro project to store them. See Section 8.1 of the *Maestro User Manual* for more information.

If you backtrack in a given run and make changes, the results for all the steps that depend on what you changed are discarded. When you make a change that affects later results, you are prompted to save the previous results and create a new run.

The step features occupy the center of the panel, and consist of a title with a brief description of the step at the top, a set of controls and tables for results, and a Back and a Next button at the bottom. The tasks for each step are summarized in the next section, and the steps are described in detail in the following chapters.

In addition to the step features, the panel contains a menu bar and an octagon button at the top, and a step guide just above the Close and Help buttons. These features are described below.

Table 5.1. Description of the File menu.

Menu Item	Description
New	Create a new run. Opens a dialog box to specify the run name. The new run becomes the current run.
Open	Open an existing run from the submenu. If there are more than 4 runs, choose More to open a dialog box and select a run.
Save As	Save the current run with a new name. Opens a dialog box to specify the new run name. The run is copied, and the renamed run becomes the current run.
Rename	Rename the current run. Opens a dialog box to specify the new run name.
Delete	Delete the current run.

The File menu

The File menu allows you to work with the runs that are available in the project. The items on this menu are described in Table 5.1.

The Step menu

The Step menu contains an item to display or hide the Guide, and items for each of the steps. The current step is marked with a red diamond. If the Guide is displayed, it is marked with a red square. The items for the steps that are not available are dimmed. You can go to any available step by choosing the corresponding menu item.

The octagon button

When a job has been launched and is running, the gray octagon at the upper right of the panel turns green and spins. To monitor the job using the Monitor panel, click the octagon. For more information about monitoring jobs, see Section 2.10 on page 28 or Chapter 5 of the *Job Control Guide*.

The Guide

The Guide displays the steps in the model as a set of buttons linked by lines. The buttons for the steps that are not available are dimmed. The current step is highlighted with a white background. You can go to any available step by clicking its button in the Guide. The Guide can be displayed or hidden from the Step menu.

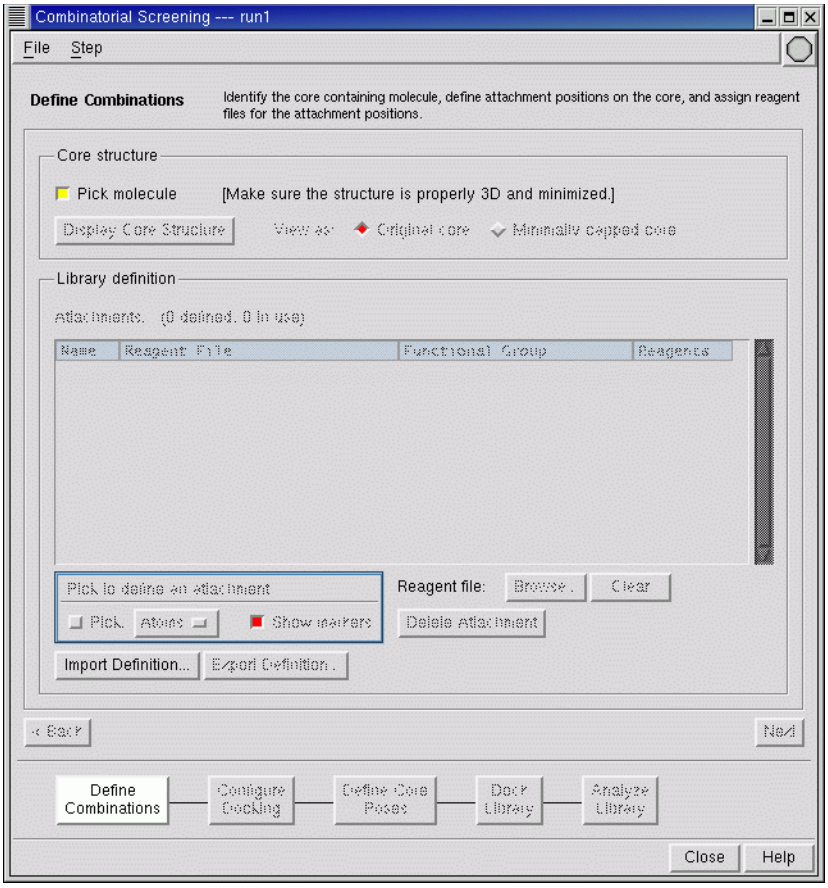


Figure 5.1. The Combinatorial Screening panel.

5.2 The Combinatorial Screening Process

The tasks involved in each step of the combinatorial screening process are summarized below.

The Define Combinations step

In this step, you first select the core-containing molecule that you want to use for generation of the library. The library is generated by replacing selected side chains in the molecule with other side chains from the reagent files. For each side chain, you identify the bond that defines where the substitution will take place (the *attachment position*), and specify a reagent file that contains the reagents with the side chains that you want to substitute. The bond is called the *grow bond*. The combination of an attachment position and a reagent file is called an *attachment*. The core of the molecule is the part of the molecule that is left when you remove the side

chains. The reagent libraries must be prepared beforehand, using the Reagent Preparation panel.

Once you have selected a molecule and the attachment positions for the reagents, and added the reagent files, you can proceed to the next step.

The Configure Docking step

The screening process uses CombiGlide docking to filter out reagents that produce molecules that do not score well. This step assumes that you have already generated the necessary Glide grid files, using the Receptor Grid Generation panel. In this step, you set options for the docking of the libraries to the receptor. As well as selecting a grid and making basic settings as for a Glide run, you can specify the full range of Glide constraints for the docked molecules.

After selecting a grid and setting any desired options, you can proceed to the next step.

The Define Core Poses step

In this step, you select the method for determining the poses of the core structure that you want to use in the docking step. These poses are generated by docking one or more molecules that contain the core. These molecules can be the core-containing molecule, the minimally capped core, or some other molecules that you choose. The *minimally capped core* is the core of the molecule with a minimal capping group at each attachment position. In most cases, the minimal capping group is defined by R, R', R"=Me or Ar=Ph for each reagent. See Table 4.1 on page 39 for definitions of R, R', and R".

You can constrain the core position to prevent it from moving to regions of the receptor that would generate unacceptable binding modes.

When you have selected a method, you can proceed to the next step.

The Dock Library step

In this step you run the docking jobs. The docking is performed in three stages. In the first stage, the core-containing molecules are docked to obtain a set of core poses, using whatever constraints on the core position you imposed.

In the second stage, each reagent from each reagent library is added in turn to the core, to produce molecules with a single side chain and the minimal capping groups at the remaining attachment positions. These molecules are docked using CombiGlide XP docking. Reagents that do not generate any good poses are eliminated. This stage can be run independently, and is known as the “single-position docking” stage.

In the third stage, the poses with a single side chain are combined intelligently to produce molecules that are likely to dock well. These molecules have side chains from the reagent

libraries at all attachment points. This stage is the most time-consuming part. Glide constraints are only applied in this stage. This stage is known as the "combinatorial docking" stage.

After docking, you can export the results of either the single-position docking or the combinatorial docking for further examination. From this step you can also enumerate and dock the entire library without performing the initial screening.

Once you have docked the combinatorial library, you can proceed to the next step.

The Analyze Library step

In this step, the number of reagents used at each position is reduced to a small set by application of filtering and selection strategies. Filtering is based on structural properties such as molecular weight and counts of structural features, and ADME properties. These properties are generated by QikProp, which is run automatically if it is installed. The selection strategies identify the "best" reagents at each position, based on the GlideScore of each of the docked structures. When the small reagent set is selected, the library defined by this set can be enumerated.

Defining the Core and Its Attachments

The combinatorial screening process begins with the definition of the core and its attachments. The core-containing structure must be a 3D, all-atom, minimized structure. Likewise, the reagents used for the attachments must be 3D, all-atom, minimized structures, and also include information that defines the side chain to be attached to the core. Preparation of the core-containing molecule and the reagents is described in Chapter 4. The task of selecting the core-containing molecule, defining the attachment points, and associating a set of reagents with each attachment point is carried out in the Define Combinations step of the Combinatorial Screening panel.

6.1 Selecting the Core-Containing Molecule

The first task in this step is to select the core-containing molecule. When you first enter the Define Combinations step, Pick molecule is selected in the Core structure section, and most other controls are unavailable (see Figure 6.1).

To select the core-containing molecule:

1. Ensure that the desired molecule is displayed in the Workspace.

   You can import the structure if necessary, or open the Project Table panel to display the structure, while the Combinatorial Screening panel is open. If you import the structure, ensure that Replace Workspace is selected.

2. Click on an atom in the core-containing molecule.

   Any structures other than the core-containing molecule are undisplayed. The other controls in the step become available.

The molecule that you picked is copied into the CombiGlide run. Any changes you make to the structure of the molecule are not saved in the run. If you make changes to the structure, you must pick the molecule again to make it the core-containing molecule. If you have already picked a molecule, picking a new molecule deletes all information about attachments.

After you pick the core-containing molecule, Pick molecule is deselected, and most other controls become available. The Pick option is selected in the Pick to define an attachment section. The core-containing molecule is now defined, and you can define the positions of the attachments, which also defines the actual core.

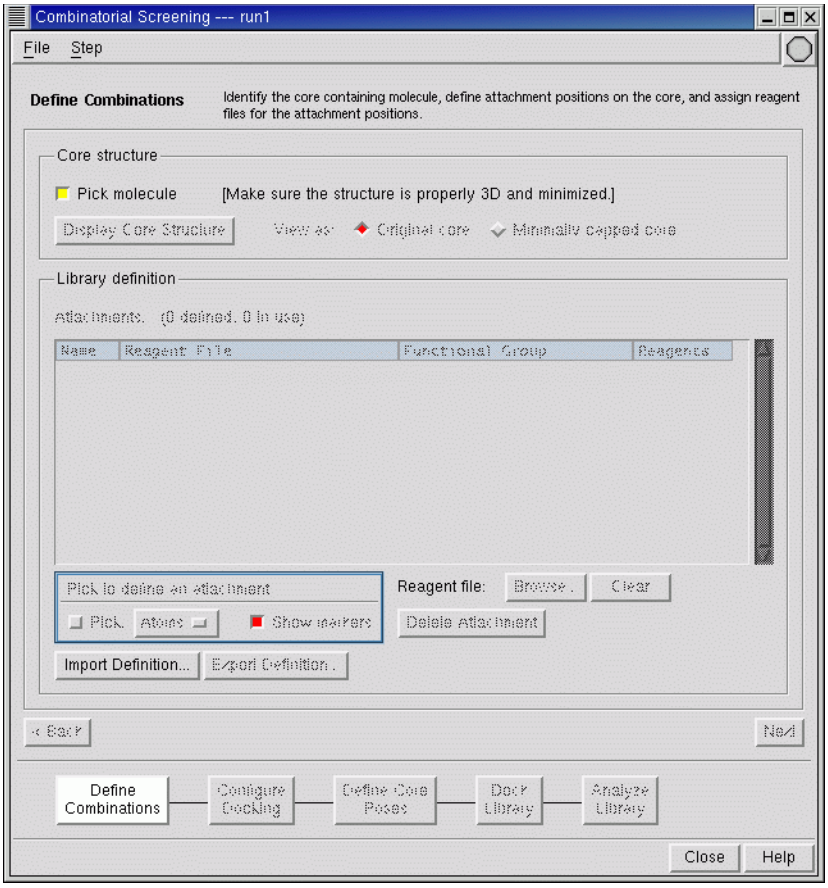


Figure 6.1. Initial view of the Define Combinations step.

6.2 Defining Attachments

Once you have a core-containing structure, you can define the attachments. An *attachment* is defined as the combination of an attachment position and a file that contains the reagents that will be attached at this position. The attachment position is defined by the atom in the core to which a side chain is attached and the atom in the side chain that is bonded to it. To make the definition, you must select both the atom that is kept in the core, and the atom that is bonded to it in the side chain. The order in which you pick these atoms determines which is the core and which is the side chain. The bond so defined is called the *grow bond:* it is essentially a vector that points from the core to the side chain. The direction is the direction in which the side chain is "grown" onto the core.

Having defined the attachment position and the side chain, the next task is to select a set of reagents for this attachment position from which the side chains are taken. The reagents are contained in files that you prepared using the Reagent Preparation panel.

To define an attachment:

1. Pick two atoms in the core-containing molecule that define the attachment position, the atom that is kept first, then the atom that is deleted.

   Alternatively, choose Bond from the Pick option menu in the Pick to define an attachment section, and click on the appropriate bond. When you click, make sure you click on the end of the bond closest to the side chain, to ensure that the grow bond points in the right direction. This is especially important for the first attachment. For subsequent attachments, a warning is posted if the grow bond points in the wrong direction.

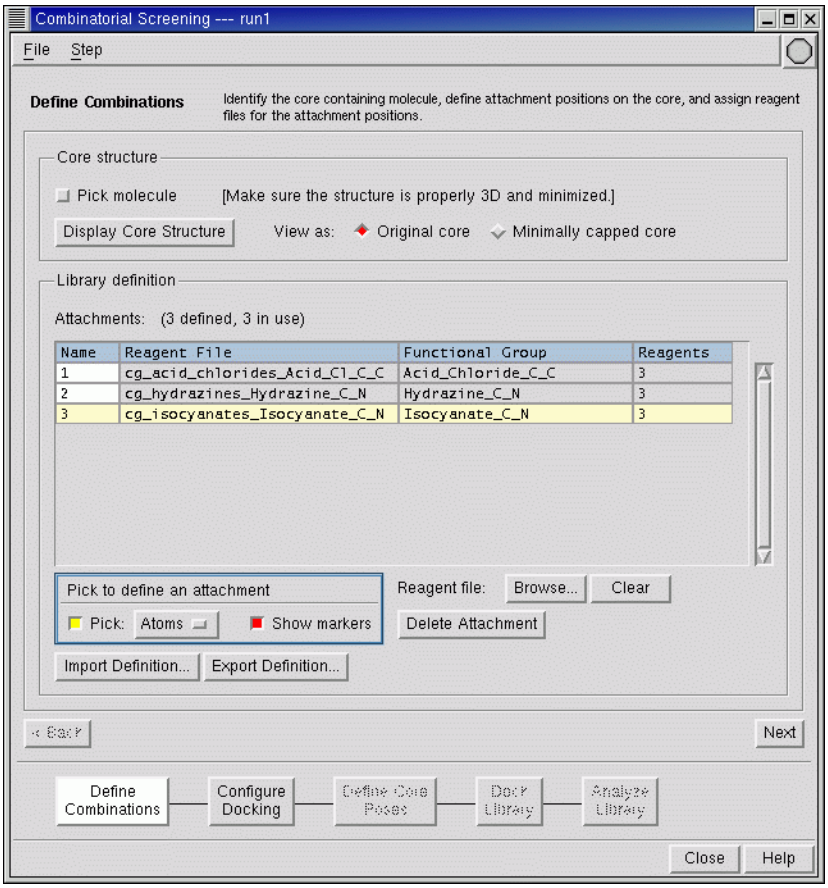


Figure 6.2. The Define Combinations step after defining attachments.

After picking the second atom (or the bond), a gold arrow is displayed over the bond, and the Select Reagent File dialog box opens.

2. Select the appropriate reagent file, and click OK.

   The reagent files have a `.bld` extension: they are Maestro-format files that have special information for the Maestro builder, which does the library enumeration. The dialog box closes, and the information in the row of the Attachments table is completed. The size of the library with the current set of attachments is given in the lower portion of the panel.

   The arrowhead is replaced with a gold sphere centered on the side chain atom.

If you choose the wrong location or the wrong direction for an attachment, you can delete it by selecting it in the table and clicking Delete Attachment. When the error is recognized by CombiGlide, such as if you pick a bond in a ring, a warning is posted and the attachment is cleared. If you simply attached the wrong file, you can select the attachment in the table, and click the Browse button below the table to select the correct file, or click Clear to remove the attachment file.

For each attachment a minimal capping group is automatically defined. This minimal capping group is usually a small instance of the side chain in which R, R', R"=Me or Ar=Ph for the chosen reagent type. Adding the minimal capping group to the core defines the *minimally capped core*. You can view the minimally capped core by selecting Minimally capped core from the View as options in the Core structure section. To redisplay the original core-containing structure, select Original core.

Information for the attachments that have been defined is listed in the Attachments table. The table columns are described in Table 6.1. You can select multiple rows in the table with the usual shift-click and control-click actions. The markers for the selected attachments are colored turquoise in the Workspace.

Table 6.1. Description of Attachments table columns

Column	Description
Name	Label for the position of the attachment. Default is 1, 2, 3, etc. Editable.
Reagent File	Name of the file containing the reagents, minus the extension. The tooltip for this table cell displays the full path to the file.
Functional Group	Long name of the functional group that describes the reagents.
# Reagents	Number of reagents in the file.

Right-clicking in a table row displays a menu with four items: Select Reagent File, Clear Reagent File, Reagent File Statistics, and Delete Attachment. The actions are applied to all selected rows. The Reagent File Statistics item displays information on the selected reagent files from the reagent preparation process, including the number of structures that were generated at each step and molecular weight data statistics. The other items perform the same actions as the Browse, Clear, and Delete Attachment buttons.

6.3 Saving and Loading Library Definitions

The combination of a core structure and its attachments defines a combinatorial library. You can save this definition for future use, by clicking Export Definition. A file browser opens, in which you can navigate to the desired location and specify a name. The library definition is saved as a gzipped tar file, and the extension `-comdef.tar.gz` is added to the name automatically. The library definition includes the core structure, which is written to the file `core-comdef.mae`, and the files for each of the attachments. These files are all added to the tar file.

If you want to use a previously-defined library definition, including the core structure and the attachments, click Import Definition, and navigate to the desired gzipped tar file. The default filter for the file selector is `-comdef.tar.gz`, so you should only see valid library definitions.

Setting Up for Docking

The center of the combinatorial screening process is the docking of the library members using CombiGlide XP docking. After you have defined the library, the next task is to set up the parameters and options for docking.

The docking of the library is performed in three stages:

- **Core docking.** In this stage, the core-containing molecules are docked to determine the best core poses. By default, the minimally capped core is used. These poses are used as the initial poses for the substituted structures that constitute the library members. Two jobs are run in this stage, called `core` and `mcore` docking.
- **Single-position docking.** In this stage, the side chains at each position are added to the core, one at a time, to generate a set of structures with side-chain substitution at a single position. The other attachment positions are capped with the minimal capping group for that position. These singly-substituted structures are then docked, and passed through a selection process that eliminates reagents that are likely to result in poor binding.
- **Combinatorial docking.** The final stage is to dock a selected number of fully substituted structures that were predicted to have the highest probability of binding to the receptor by the selection algorithm.

Setting up the docking calculations is done in two steps: a configuration step, in which Glide parameters and constraints are set, and a core pose definition step, in which the structures to use for the core docking are selected and any constraints on the core position are set. These tasks are performed in the Configure Docking and Define Core Poses steps of the Combinatorial Screening panel.

7.1 Making Glide Settings

The Glide settings are made in the Settings folder of the Configure Docking step. This folder is a combination of parts of the Settings folder and the Ligand folder from the Glide Ligand Docking panel. Because CombiGlide uses Glide XP docking and the ligands are predefined, only the relevant options from these folders are available. There are three sections in this folder: Receptor grid, Docking, and Van der Waals radii scaling.

The first task is to select the receptor grid. You can either enter the path to the receptor grid file in the Receptor grid base name text box, or click Browse and navigate to the receptor grid file.

If you have not generated a grid, you must do so before proceeding, in the Receptor Grid Generation panel. You can open this panel from the CombiGlide submenu of the Applications menu in the main window. When you have generated a grid, you can reopen the Combinatorial Screening panel and return to this step.

Next, you can set options and parameters as listed below. The ring flip and amide bond rotation options are applied in all docking stages. The limits on the maximum number of atoms and rotatable bonds are only applied in the combinatorial docking stage. Structures exceeding these limits are not built and thus are not docked. For more information on these options, see Chapter 7 of the *Glide User Manual*.

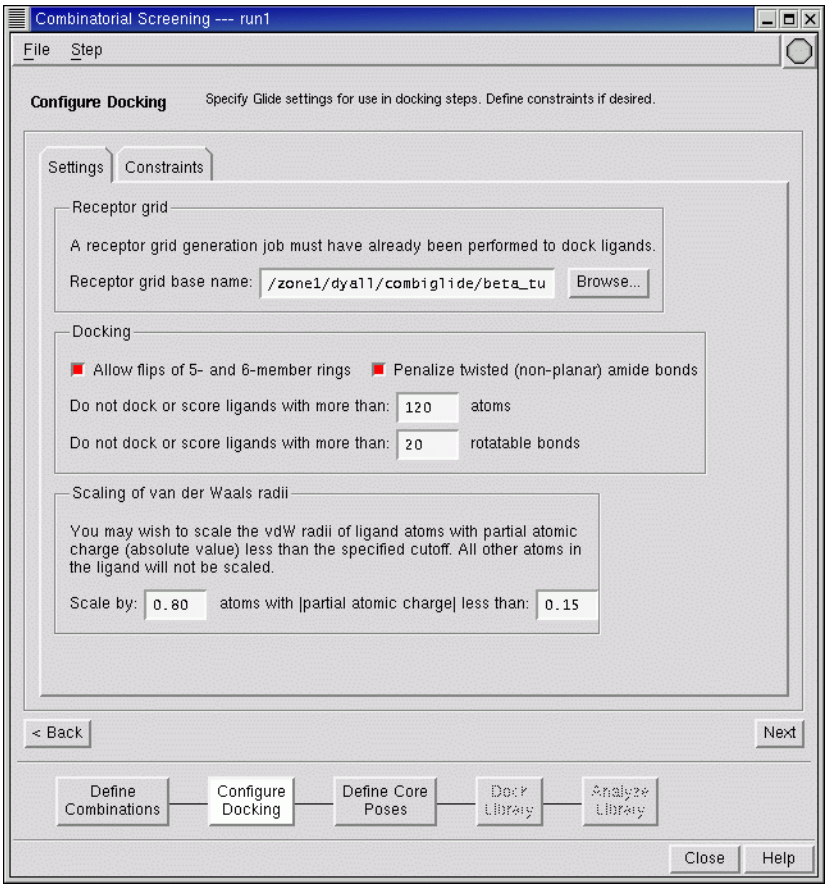


Figure 7.1. The Settings folder of the Configure Docking step.

Allow flips of 5- and 6-member rings

Select this option to allow the conformation of nonaromatic 5- and 6-membered side-chain rings to be varied during docking. This option is selected by default.

Penalize amide bond rotations

Select this option to penalize rotations around amide C-N bonds when scoring the docked molecules. This option is selected by default.

Do not dock ligands with more than *n* atoms

Enter a number to restrict the size of the molecules that are docked. The default is 80. This restriction is only applied in the combinatorial docking stage.

Do not dock ligands with more than *n* rotatable bonds

Enter a number to restrict the number of rotatable bonds to sample during docking. The default is 15. This restriction is only applied in the combinatorial docking stage.

Van der Waals radii scaling

Enter a scaling factor and a partial charge in the text boxes to scale the van der Waals radii of atoms with small partial charges. The scaling simulates the flexibility of nonpolar groups.

7.2 Setting Glide Constraints

Constraints can also be used with CombiGlide and are specified in the same manner as with Glide. With CombiGlide, constraints are only turned on during the combinatorial docking stage, i.e. docking of the fully substituted structures. This prevents the docking algorithm from forcing the side chains at different positions on the core into the same binding pocket during the single-position docking. In the combinatorial docking stage, constraints can be used to filter out reagents whose side chains cannot satisfy the relevant constraints.

The Constraints folder of the Configure Docking step folder contains controls for applying Glide constraints to the docking of the combinatorial library. The contents of this folder are identical to those of the Glide ligand docking Constraints folder. For information on setting constraints, see Section 7.4 of the *Glide User Manual*.

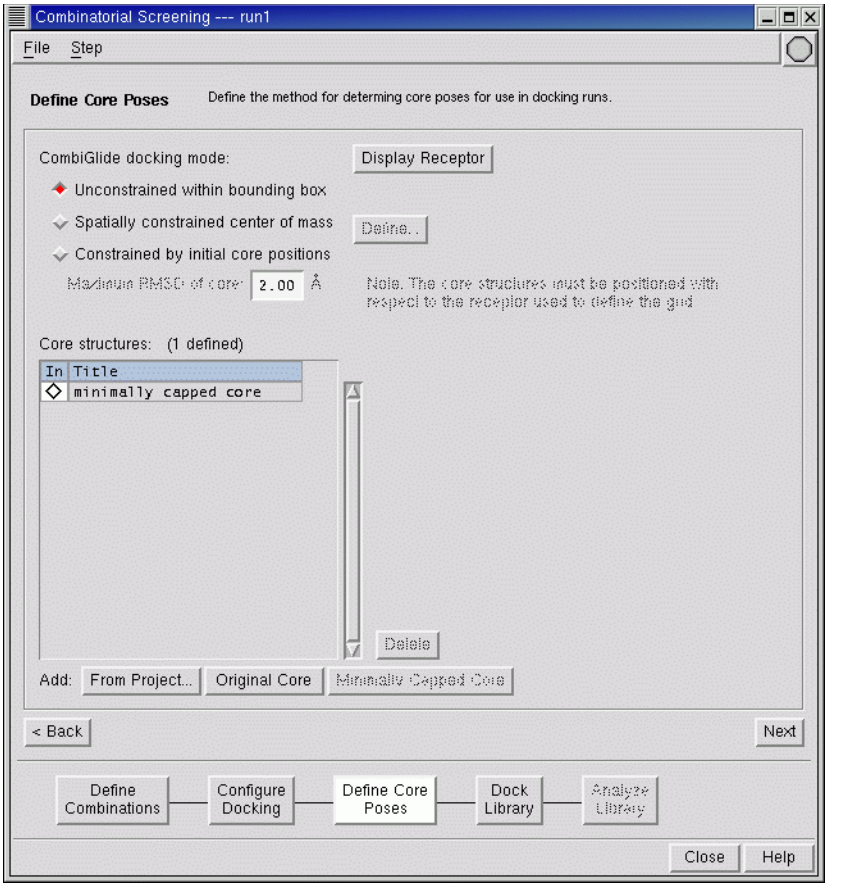


Figure 7.2. The Define Core Poses step.

7.3 Defining the Core Poses

The final setup task is to determine how the core poses will be generated in the first stage of the docking step. The core poses can be generated with any molecule that includes the core, not just the structure you selected to define the core in the Define Combinations step. This task is carried out in the Define Core Poses step, and involves selecting a docking mode and specifying the structures that will be docked.

7.3.1 Selecting a Docking Mode

The first stage in the docking step is to dock the core-containing molecules. This stage does not apply any Glide constraints that you have set. If the molecules used to define the core poses are

smaller than the original core-containing molecule, they might bind to locations in the active site that are not accessible to the original. Many of these poses could be undesirable, for example if the core migrated to a location in which the fully substituted molecules could not dock. CombiGlide therefore supplies three docking modes, described below, two of which allow some kinds of constraints to be applied.

Unconstrained within bounding box

This is the default option and permits the molecule the widest range of movement within the grid bounding box during docking. If the molecules you select for the core poses are likely to dock in reasonable poses, select this option.

Spatially constrained center of mass

With this option, the core center of mass is constrained during docking to a sphere of a given radius. Constraining the center of mass allows the core-containing structures to rotate but prevents them from moving to another possible binding site.

The position of the sphere is defined by picking atoms in the Workspace, and you can choose atoms in the ligand or the receptor, or both. If you want to display the receptor in order to pick atoms, click Display Receptor. You can also display any of the core-containing structures by clicking the appropriate In column of the Core structures table. However, you should make sure that these structures are aligned to the receptor, otherwise picking atoms from them will be of no value. When you have the desired structure in the Workspace click Define. In the Define Core Constraint dialog box, use the picking tools to pick the atoms whose centroid defines the sphere center.

The radius of the sphere is 1.0 Å by default, and can be set in the Define Core Constraint dialog box. The sphere radius must not extend beyond the enclosing box of the grid. You can also set the coordinates of the center of mass manually.

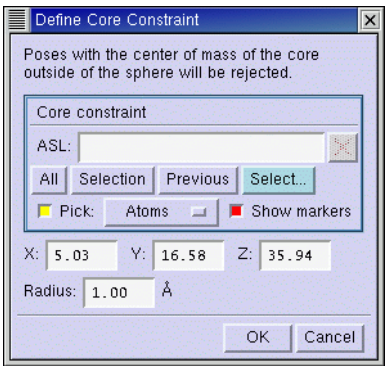


Figure 7.3. The Define Core Constraint dialog box.

Constrained by initial core positions

This option allows you to define the core poses by the positions of the core-containing structures, and provides the greatest degree of constraint on the core poses. These structures must be already aligned to the receptor. When these structures are docked, poses for which the RMSD of the core atoms exceeds the specified threshold are rejected. One way of obtaining these core-containing molecules is to dock molecules that contain the core onto the same receptor using Glide, with whatever constraints you want to apply. Another is to use the crystal structure of the chosen receptor complexed with a core-containing molecule.

7.3.2 Specifying Structures for the Core Poses

The structures that will be docked to define the core poses are listed in the Core Structures table. You can add structures to the table with the Add buttons below the table, and you can delete structures from the table by selecting them and clicking Delete.

By default, only the minimally capped core is listed. To add the original core-containing molecule, click Original Core. To add structures from the Project table, click From Project, and select the molecules in the entry chooser that opens. The entries must satisfy the following criteria:

- They must contain only a single molecule
- The molecule must contain the core structure

No checking is done for duplicates, so you must ensure that you don't add the same structure more than once. The structures are copied into the run, so that any changes in the Project Table do not affect these structures.

If a crystal structure or docked pose is used for the core structure, the associated protein structure must be superimposed on the receptor structure from which the grids were calculated to ensure that the core poses are in the correct frame of reference. You can superimpose the protein structures in the Superposition panel, which you open from the Tools menu in the main window.

Docking the Library

The docking step is the heart of the combinatorial screening process, where the docking jobs are run to screen out molecules that do not dock well and thereby eliminate reagents. This step is carried out in the Dock Library step of the Combinatorial Screening panel.

First, the core structures are docked. The resulting poses are clustered, and poses from each cluster are selected until the maximum number is reached. This ensures the widest diversity of core poses while limiting their number. The core poses are used as the initial positions of the substituted molecules in the subsequent docking stages.

The screening is performed in two stages. In the first stage, structures that have only a single substitution are docked. These structures are generated by adding the reagents, one at a time, to the minimally capped core. The result is a set of molecules (a "single-position library") for each attachment position with a side chain at that position and the minimal capping group at the other positions. The side chains added are taken from the expanded set of reagents, in which different ionization states, stereoisomers, conformers, or tautomers can be represented. In the pyrazole example, there are three attachment positions and thus, three single-position libraries will be generated and docked. This portion of the virtual library evaluation is an additive process: the number of dockings required for this step is the sum of the number of side chains for each position.

Once the single-position docking results are returned, a selection algorithm determines which fully substituted cores have the highest probability of binding to the receptor. Structures from the single-position docking stage that do not dock well are eliminated. Reagents for which no structures docked well are removed from the reagent list for the appropriate attachment position. The selection algorithm takes into consideration the XP GlideScores of the docked poses from the single-position docking. It also checks that the side chains it has identified for a particular fully substituted core are not predicted to compete for the same part of the receptor binding site.

Finally, the structures that are identified by the selection algorithm are generated and docked. This set of structures is not a combinatorial library, in which each side chain occurs at each position.

As an alternative, the single-position docking and the selection procedure can be bypassed, and the entire library can be enumerated and docked.

8.1 Running the Docking Jobs

The two stages of docking—the single-position docking and the combinatorial docking—can be run in the same job, or run separately. If you want to examine the results of the single-position docking before continuing with the combinatorial docking, click Single Positions. When you want to proceed to the second stage, click Combinatorial. The results of the first stage are used to run the combinatorial docking. To run both stages in the same job, click Combinatorial.

In both cases, the core docking is run first, then the single-position docking. For the combinatorial docking, you can set a limit on the number of fully substituted structures to dock in the Maximum number of combinatorial structures to dock text box. The default is 20 times the sum of the number of side chains at each position. If you plan to use Glide constraints, you should significantly increase the limit to ensure sufficient output structures.

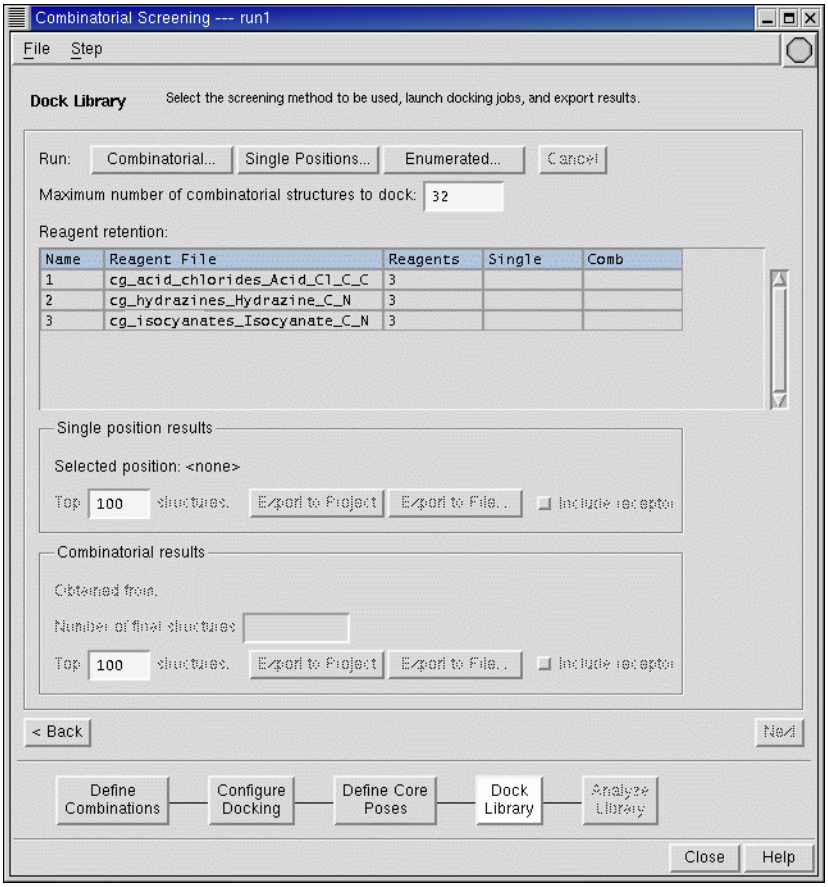


Figure 8.1. The Dock Library step.

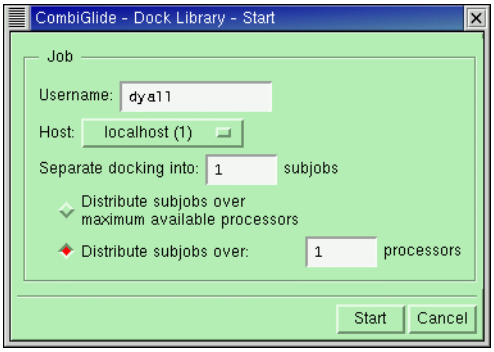


Figure 8.2. The Start dialog box.

When the final docking stage finishes, if QikProp is available, it is run for the completed structures, and the QikProp properties are added to the run. These properties can be used in the Analyze Library step.

If you want to dock the entire combinatorial library, click Enumerated. A dialog box is displayed, in which you can select the Glide docking mode. In addition to using CombiGlide XP docking, you can select any of the Glide docking modes: XP, SP, or HTVS. You must have the appropriate Glide license for Glide docking. The limit in the Maximum number of combinatorial structures to dock text box does not apply in this case: the entire library is docked.

When you click Single Position or Combinatorial, or click OK in the options dialog box for enumerated docking, a Start dialog box is displayed, in which you can make job settings. These include selection of a host, and distribution of the job over multiple processors.

To cancel the current docking job, click Cancel, then click OK in the confirmation dialog box. If you are running the combinatorial docking job, and the single-position docking job has already finished, the results of the single-position docking job are retained.

8.2 Docking Results

Statistics on the reagents that survive the various docking stages are added to the Reagent retention table when the docking jobs are finished. The table is noneditable, and its columns are described in Table 8.1. The number of structures actually generated is reported in the Combinatorial results section. This number can be less than the product of the numbers in the # Comb column of the Reagent retention table, because of the limits imposed and because not all structures generated from the single-position docking necessarily dock well.

Once the results are available, you can export them to the Project Table or to a file. The structures from both stages are sorted by their GlideScores. Because there may be variations of each

Table 8.1. Description of Reagent retention table.

Column	Description
Name	Attachment position label
Reagent File	Name (minus extension) of the file that contains the reagents for the specified position.
#input	Number of reagents in the original reagent file.
#single	Number of reagents from this file that survived the single-position docking.
#comb	Number of reagents from this file that survived the combinatorial docking. This number includes all reagents that contributed to any structure that docked well.

reagent (such as different ionization states, tautomers, conformers), it is the top-scoring variation for each reagent that is exported. For the single-position docking results, multiple poses are included in the set that is exported: thus the top-scoring variation for each pose for each reagent is exported. The set of singly substituted or the set of fully substituted structures is exported to the Project Table as an entry group, with QikProp properties if these are available. If the receptor is included, it is exported as the first entry in the entry group or the file.

To export a set of structures from single-position docking:

1. Select the row in the Reagent retention table for the desired reagent type.
2. Enter the number of reagents in the Top *n* reagents text box in the Single position results section.
3. (Optional) If you want to include the receptor, select Include receptor.
4. Click Export to Project or Export to File.
5. If you are exporting to a file, navigate to the directory, enter the file name, and click Export.

To export a set of structures from combinatorial docking:

1. Enter the number of structures in the Top *n* structures text box of the Combinatorial results section.
2. (Optional) If you want to include the receptor, select Include receptor.
3. Click Export to Project or Export to File.
4. If you are exporting to a file, navigate to the directory, enter the file name, and click Export.

Focusing the Library

The final step in the combinatorial screening process is to focus the library to a small number of compounds generated from a few reagents at each position. This focused library should contain a large proportion of compounds that are likely to bind well. The focusing is carried out by applying various selection strategies, by filtering on molecular and predicted ADME properties, by manual selection, and by setting limits on the number of reagents at each position and on the overall library size. The purpose of this step is to select sets of optimal reagents for each attachment position that are small enough and chosen well enough for practical synthesis and screening of the resulting real combinatorial library.

The selection and filtering can be repeated as many times as you like, and the results of each selection and filtering job can be stored and compared. You can combine selection, filtering and manual inclusion or exclusion of reagents. When you are satisfied with a selection, you can run a job to create and save the structures in the library.

The focusing of the library is driven from the Analyze Library step of the Combinatorial Screening panel. The selection and filtering tasks are set up and run from the Filter and Select dialog box, which you open by clicking Filter and Select. This dialog box is nonmodal, so you can interact with the Workspace and the Combinatorial Screening panel while it is open.

9.1 Selection Strategies

CombiGlide has three automated selection strategies, which can be supplemented with manual selection. The automated strategies use the GlideScore to select the reagents that produce the greatest number of structures that have a good GlideScore.

The selection strategies work by filling lists of reagents at each position, until the limits on the minimum and maximum number of reagents and the overall library size are met. The molecules from the docking step are ordered according to their GlideScore or some derived quantity. Each molecule is considered in turn, and the reagents that were used to build that molecule are added to the corresponding lists.

The strategies are selected in the Selection Strategy folder of the Filter and Select dialog box.

The first strategy selects reagents based on the single-position docking. In this strategy, the lists for each position can be filled independently, from the docking results for each position. This is the only strategy available if you do not have combinatorial docking results. To apply this strategy, select Single position under Selection strategy, and click Apply or OK.

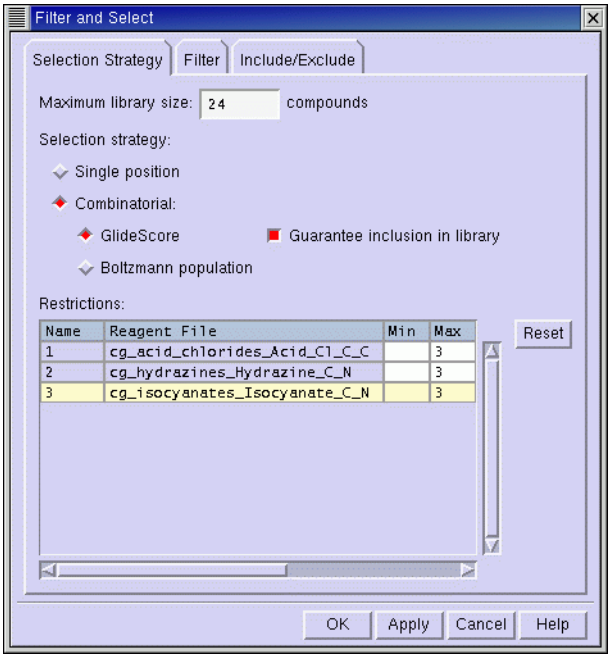


Figure 9.1. The Selection Strategy folder of the Filter and Select dialog box.

The second strategy selects reagents based on the GlideScores from the combinatorial docking. Starting from the structure with the best GlideScore, the reagent for each position is added to the list for that position. The reagents for the next best structure are then added to the lists, unless they have already been added, and so on, until the conditions on the minimum and maximum reagents at each position and on the maximum library size are met. When a list at a particular position is filled, there are two options for adding structures to the remaining partially filled lists. The first option is to add reagents from only those molecules that have reagents at the filled position that are in the filled list. The second option is to add reagents from any molecule, regardless of whether the reagent at the filled position is in the filled list. The first option ensures that all molecules used to determine the library are in the library. The second option ensures that the reagents came from the structures with the best GlideScores. To apply this strategy, select Combinatorial under Selection strategy, then select GlideScore. If you want to ensure that the library contains the molecules used to determine it, select Guarantee inclusion in library. This option is only available if you have minimum and maximum numbers in the Restrictions table (see below). Click Apply or OK to start the selection job.

The third strategy uses the GlideScore from the combinatorial docking as a free energy, and calculates the normalized Boltzmann population for each docked molecule. The populations are summed for each reagent, and the reagent populations are converted to a relative

GlideScore using a free-energy relation. Reagents are added to the lists in order of relative GlideScore, from lowest to highest. This strategy samples the most frequently occurring reagents, weighted according to GlideScore: reagents coming from molecules with good GlideScores are weighted more heavily than those with poor GlideScores. For each position, the list of reagents is filled, starting from the reagent with the lowest relative GlideScore, until the conditions on the minimum and maximum reagents at each position and on the maximum library size are met. To apply this strategy, select Combinatorial under Selection strategy, then select Boltzmann population. Click Apply or OK to start the selection job.

In addition to applying a selection strategy, you can set the minimum and maximum number of reagents at each position, and set an overall library size. The maximum library size can be entered in the Maximum library size text box. The minimum and maximum number of reagents at each position can be set by editing the table cells in the Restrictions table, which is described in Table 9.1.

Table 9.1. Description of the Restrictions table.

Column	Description
Name	Attachment position label
Reagent File	Name of reagents file for this attachment position, without the extension.
Min	Minimum number of reagents to include at this position. Editable.
Max	Maximum number of reagents to include at this position. Editable.

9.2 Manual Selection of Reagents

You can manually select reagents that must be included in the library or that must be excluded from the library in the Include/Exclude folder of the Filter and Select dialog box. This allows you to enforce the selection of reagents that did not score well, or exclude reagents that did score well, based on your knowledge of the chemistry, or to exclude reagents based on their availability.

To select reagents for inclusion or exclusion:

1. Select an attachment position from the Position option menu.

   All the reagents in the reagent file for that position are displayed in the Reagents list.

2. (Optional) Filter the reagents by entering a string in the Filter text box.

   You can use the "*" character to represent an arbitrary string. All reagents whose titles contain the specified string are listed in the Reagents list when you press ENTER.

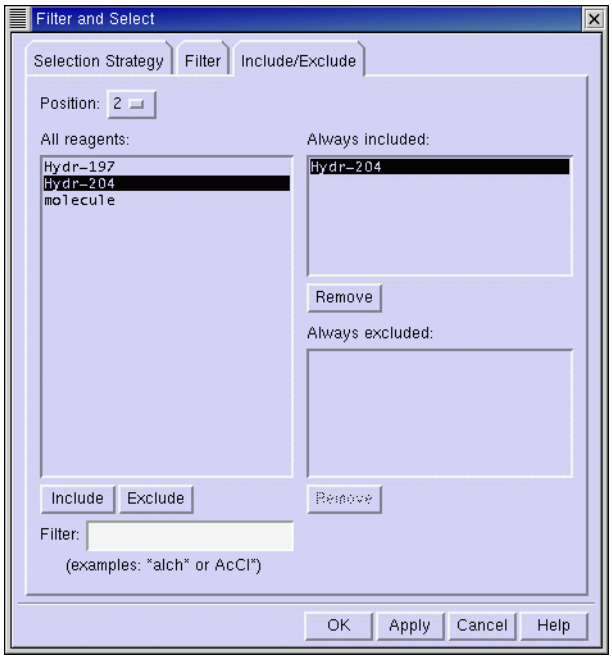


Figure 9.2. The Include/Exclude folder of the Filter and Select dialog box.

3. Select reagents from the Reagents list.
4. Click Include to add the selected reagents to the Always Included list, or click Exclude to add the selected reagents to the Always Excluded list.

If you add a reagent to one list that is already in another list, it is removed from the other list.

To remove reagents from a list:

1. Select the reagents in the list.
2. Click the Remove button immediately below the list.

You can include or exclude reagents directly from the Analyze Library step, by right-clicking a reagent in the Reagents at position *N* table. When you do so, the Filter and Select dialog box opens so that you can re-run the library selection job with the newly included or excluded reagent.

9.3 Filtering

In addition to selecting reagents with one of the selection strategies, you can filter the combinatorial results based on various predicted ADME properties of the molecules. These properties are generated by QikProp. Several filters are provided, and you can customize these filters by including or excluding any of the listed properties and by setting the limits on the acceptable range of the properties. The property ranges in the filters represent the undesirable property ranges: molecules whose properties fall inside these ranges will be filtered out. The filters are set up in the Filter folder of the Filter and Select dialog box.

If you apply both a filter and a selection strategy, the filtering is performed first, then the selection strategy is applied to the results of the filtering. Manual selection overrides both filtering and the selection strategy.

The Filter set option menu provides three preset filters, Druglike, Leadlike, and Coarse, a customizable filter (Custom), and an option to bypass filtering (None). If you modify one of the preset filters, the menu item switches to Custom.

Filtering is not available unless you have a QikProp license, and have combinatorial docking results.

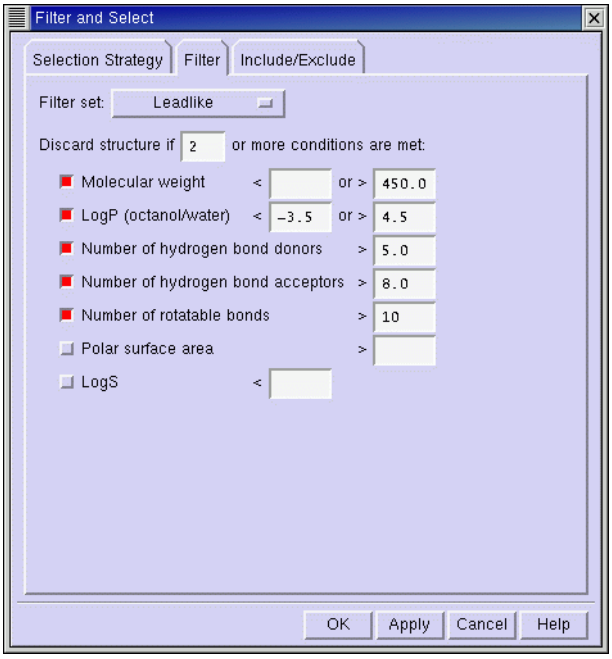


Figure 9.3. The Filter folder of the Filter and Select dialog box.

To use a predefined filter:

1. Choose a filter from the Filter set option menu.

To use a custom filter:

1. Choose a filter set from the Filter set option menu.

   Choose Custom for a blank filter form, or choose one of the predefined filter sets to alter.

2. Enter the number of conditions that must be met for the molecule to be discarded in the Discard structure if *N* or more conditions are met text box.

3. Select the properties that you want to include in the filter set.

4. Enter the values in the text box that define the property range.

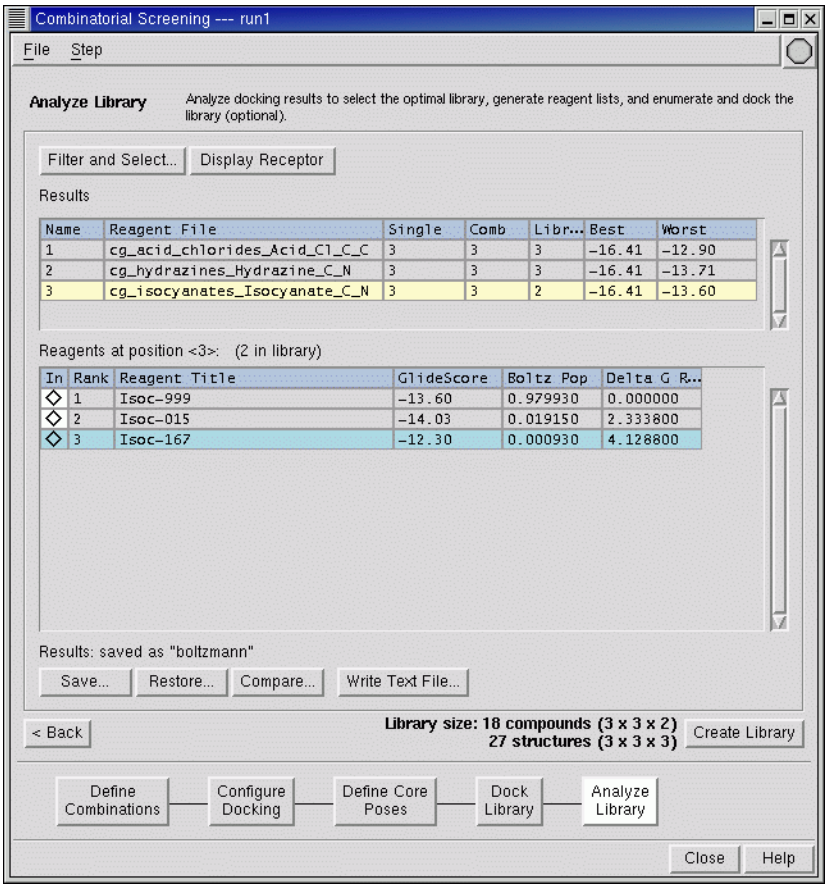


Figure 9.4. The Analyze Library step.

9.4 Results of Selection and Filtering

The selection and filtering job is run when you click OK or Apply in the Filter and Select dialog box. After the job finishes, the results of the selection and filtering process are displayed in the table in the upper part of the panel. The results—the current filter and selection settings and the results from these settings—constitute a library selection. The library size is displayed next to the Create Library button. The first four columns of the table are identical to the corresponding columns of the Reagent retention table from the Dock Library step. The columns are described in Table 9.2.

Table 9.2. Description of the results table.

Column	Description
Name	Attachment position label
Reagent File	Name of the reagent file (minus the extension).
# Single	Number of reagents from this file that survived the single-position docking.
# Comb	Number of reagents from this file that survived the combinatorial docking. This number includes all reagents that contributed to any structure that docked well.
# Lib	Number of reagents selected for the library.
Best	Best GlideScore of the molecules used to select reagents at this position for the library. If combinatorial results are used in the selection, the best GlideScore is the same for all reagents.
Worst	Worst GlideScore of the molecules used to select reagents at this position for the library.

When you select a row in the results table, the reagent list with results for that position is displayed in the Reagents at position *N* table. This table displays the top reagents for the position selected in the results table along with various properties (see Table 9.3). The list includes both the reagents that were chosen for the library at this position and a number of the best reagents that were not chosen. The reagents that were not chosen for the library are highlighted in blue.

In addition to examining the numerical results of the selection and filtering, you can view the best-scoring structure for each reagent in the Workspace. To do so, select the desired position in the results table, then click the In column of the Reagents at position *N* table for the reagent that you are interested in. The best-scoring singly substituted or fully substituted structure that contains the reagent is displayed. If you want to view the structure with the receptor, click Display Receptor. To remove the receptor, clear the Workspace.

If you decide that you want to include or exclude a reagent at a particular position, right-click in the row of the Reagents at position *N* table for the reagent and select Add to Include List or Add to Exclude List from the shortcut menu that is displayed. Selecting a menu item opens the Filter and Select dialog box so that you can re-run the library selection job with the newly included or excluded reagent.

Table 9.3. Description of the Reagents at position table.

Column	Description
In	Workspace inclusion status. You can use this column to display the best-scoring molecule for that reagent in the Workspace. However, manually included reagents that did not contribute to any docked molecule in the docking stage cannot be included in the Workspace.
Rank	Numerical rank of the reagent. The reagents are listed in order of rank; the rank is determined by the lowest GlideScore of any molecule that contains this reagent. A + sign indicates that the reagent was manually selected for inclusion in the focused library. Manually included reagents are listed at the top and are assigned the top ranks.
Reagent Title	Title of the reagent. This is the title chosen in the Reagent Preparation panel.
GlideScore	Best GlideScore for any compound in the library that used this reagent.
Boltz Pop	Boltzmann population for the reagent. Boltzmann population strategy only.
Delta G Rel	Effective relative GlideScore for this reagent treated as a free energy, derived from the Boltzmann populations. The best reagent has a Delta G Rel of zero. Boltzmann population strategy only.

Running a new selection and filtering job overwrites the results of the current job (the current library selection). You can save the results in the run, by clicking Save. The Save Results dialog box is displayed, in which you can name the library selection. Once you have saved a library selection, you can restore it later as the current selection by clicking Restore, then selecting the library selection in the Restore Results dialog box and clicking OK. You can also delete a library selection in the Restore Results dialog box by selecting it and clicking Delete. The name of the current library selection is given immediately above the Save button—or if it is not saved, a message to indicate that has not been saved.

Saved library selections are stored within the run, and are not available outside the project. To save a library selection to an external file, click Write Text File, and navigate to the desired location in the file selector that is displayed. The contents of the file include the options used to run the library selection job, and the Reagents at position *N* table for each attachment position.

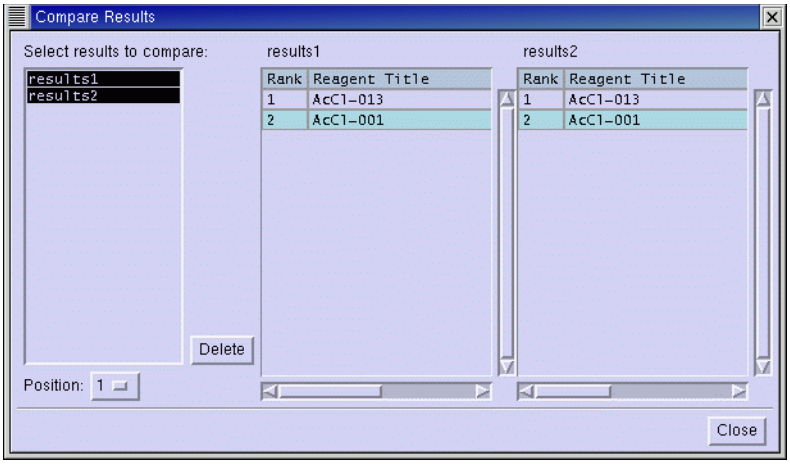


Figure 9.5. The Compare Results dialog box.

9.5 Comparing Results

In the process of selecting the optimal library, you may want to compare several library selections. To do so, these definitions must be saved in the run, using the Save button. To compare the saved library selections, click Compare, then in the Compare Results dialog box, select the library selections in the Select results to compare table. You can select and deselect items in this table with the usual click, shift-click and control-click actions. As you select the item, the Reagents at position *N* table for that library selection and the chosen position is displayed to the right of the list. To change the position for which results are displayed, choose the position from the Position menu. All the tables are updated with the results for the new position. You can compare up to six library selections.

You can also delete library selections, by selecting them in the Select results to compare list and clicking Delete. The selections are deleted without further confirmation, so check your choices carefully.

9.6 Creating a Library

When you have decided on the library selection, you can proceed to create the library. The library is a Maestro file that contains structures generated from the reagent variants, subject to a GlideScore threshold. A reagent variant is discarded if the GlideScore of the best-scoring molecule to which it contributed is above the threshold. The threshold is the highest GlideScore of any molecule that contributed a reagent to the library. In addition, a plain text

file is written that contains a summary of the library selection. This is the same file as is written with the Write Text File button. The library can be created with or without docking the results.

If the library is docked, the output Maestro file includes the receptor and the docked structures. The library can be docked with any of the Glide docking modes, but you must have a separate Glide license.

If the library is enumerated only, you can choose to untangle and minimize the structures. When the structures are built, it is possible that the structures are "tangled"—for example, a chain might go through the center of a ring. In CombiGlide you can choose to untangle these structures, and then minimize them. Untangling and minimizing the structures takes longer, but the results are superior. Untangling is performed by default in the Dock Library step when the fully substituted structures are created.

To create the library:

1. Click Create Library.

   The Create Library - Options dialog box opens.

2. Choose a mode for creating the library.

   You can select Enumerate only, do not dock, or Enumerate and dock with.

3. Select options for the library creation mode.

4. Click OK.

   A Start dialog box opens.

5. Select job options, and click Start.

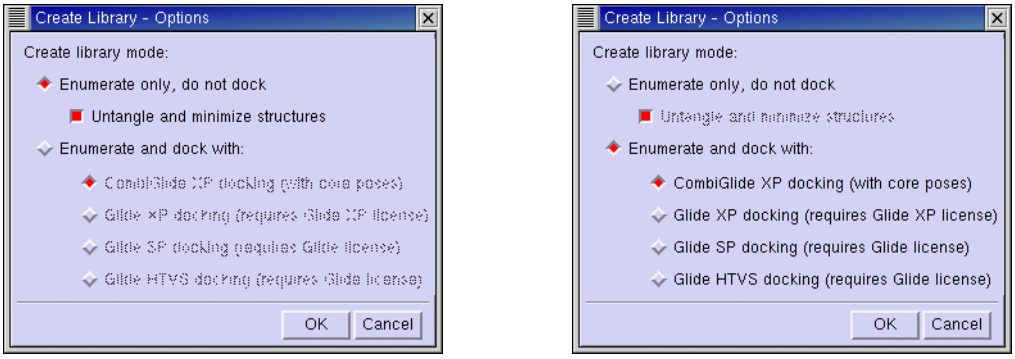


Figure 9.6. The Create Library - Options dialog box showing options for the two modes.

Enumerating a Combinatorial Library

CombiGlide provides tools for simple library enumeration in addition to the tools for focused library design, that is, for creating a library of ligands from a core structure to which side chains are added in user-specified positions. This task is carried out in the Combinatorial Library Enumeration panel, which you open from the CombiGlide submenu of the Applications menu in the main window.

The combinatorial library enumeration is carried out by substitution of side chains on the core structure with fragments from the reagent structures. For both the core and the reagents, the bonds to be broken must be designated. The fragments from the reagents are then attached to the core at the site of the broken bond. This process can be regarded as "growing" the new side chain onto the core in place of the old side chain, and the bond that is replaced (with its attached atoms) is called the grow bond. These bonds are not necessarily the bonds that are broken and formed in the real chemical reaction, but represent a means of varying the side chains on a core structure that might include part of the real reagent.

To generate a combinatorial library, you must have a 3D, minimized structure for the core-containing molecule whose side chains you want to replace, and a set of prepared reagent files that define the replacements. For information on preparing the core structure and the reagent files, see Chapter 4.

The Combinatorial Library Enumeration panel (Figure 10.1) is essentially the same as the Define Combinations step of the Combinatorial Screening panel, which is described in Chapter 6. The procedure is summarized here.

When you first open the Combinatorial Library Enumeration panel, Pick molecule is selected in the Core structure section, and most other controls are unavailable. After you have selected a core-containing molecule, the other controls become available.

To define the core-containing structure:

1. Ensure that the structure is displayed in the Workspace.
2. Click on an atom in the core-containing molecule.

   The core-containing molecule is now defined, and you can define the attachments. Pick molecule is deselected, and the Pick option is selected in the Pick to define an attachment section.

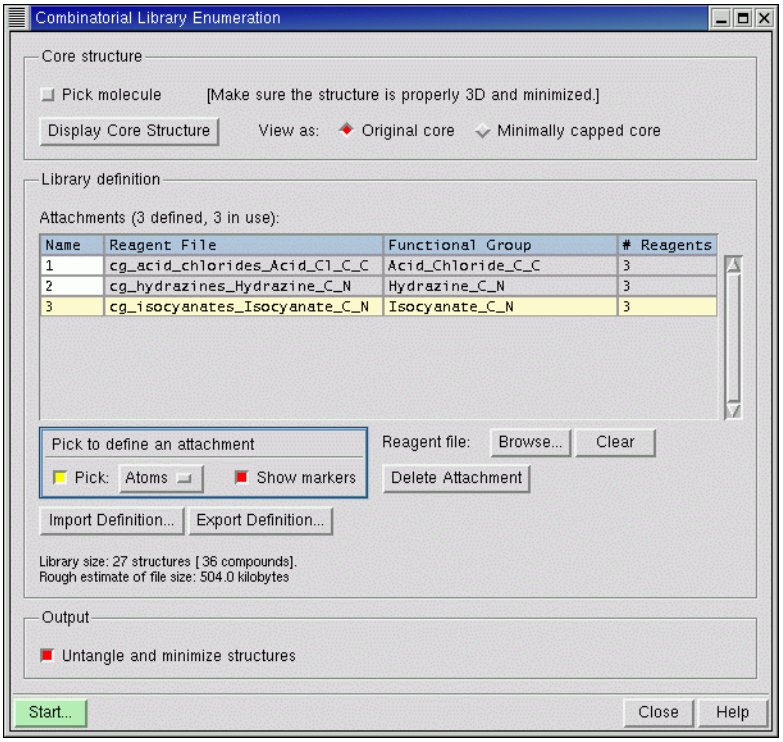


Figure 10.1. The Combinatorial Library Enumeration panel.

To define an attachment:

1. Pick two atoms that define the position of the attachment, the atom that is kept first, then the atom that is deleted.

   A turquoise arrow is displayed over the bond, and the Select Reagent File dialog box opens.

2. Select the appropriate reagent file, and click OK.

   The reagent files have a `.bld` extension: they are Maestro-format files that have special information for the Maestro builder, which does the library enumeration. The dialog box closes, and details of the attachment are listed in the Attachments table. The arrowhead changes to a sphere and the color changes to gold. The size of the library with the current set of attachments is given in the lower portion of the panel.

If you choose the wrong location for an attachment position, you can delete it by selecting it in the table and clicking Delete Attachment. If you attach the wrong file, you can select the attachment position in the table, and click Browse (below the table) to select the correct file.

When you have defined an attachment, the size of the resultant library is displayed at the foot of the Library definition section.

It is recommended that you leave the Untangle and minimize structures option selected. Untangling ensures that (for example) any chains that pass through the middle of rings during the build process are removed from the rings and the structures are properly minimized. However, if you are not interested in having accurate structures, you can deselect this option.

To start the library enumeration job:

1. Click Start.

   The Start dialog box is displayed.

2. Select job options, and click Start.

The host you choose should have temporary storage space of about 3 to 4 times the size of the library. This space is required for the untangling and minimization of the library members.

The job takes about 45-60 seconds per output structure when each structure is minimized and untangled. The library is written to a Maestro file named *jobname*`.mae`.

Getting Help

Schrödinger software is distributed with documentation in PDF format. If the documentation is not installed in `$SCHRODINGER/docs` on a computer that you have access to, you should install it or ask your system administrator to install it.

For help installing and setting up licenses for Schrödinger software and installing documentation, see the *Installation Guide*. For information on running jobs, see the *Job Control Guide*.

Maestro has automatic, context-sensitive help (Auto-Help and Balloon Help, or tooltips), and an online help system. To get help, follow the steps below.

- Check the Auto-Help text box, which is located at the foot of the main window. If help is available for the task you are performing, it is automatically displayed there. Auto-Help contains a single line of information. For more detailed information, use the online help.
- If you want information about a GUI element, such as a button or option, there may be Balloon Help for the item. Pause the cursor over the element. If the Balloon Help does not appear, check that Show Balloon Help is selected in the Help menu of the main window. If there is Balloon Help for the element, it appears within a few seconds.
- For information about a panel or the folder that is displayed in a panel, click the Help button in the panel. The Help panel is opened and a relevant help topic is displayed.
- For other information in the online help, open the Help panel and locate the topic by searching or by category. You can open the Help panel by choosing Help from the Help menu on the main menu bar or by pressing CTRL+H.

  To view a list of all available CombiGlide–related help topics, choose CombiGlide from the Categories menu in the Categories tab. Double-click a topic title to view the topic.

If you do not find the information you need in the Maestro help system, check the following sources:

- *Maestro User Manual*, for detailed information on using Maestro
- *Maestro Command Reference Manual*, for information on Maestro commands
- *Glide User Manual*, for detailed information on using Glide
- *Glide Quick Start Guide*, for Glide tutorials
- *CombiGlide Quick Start Guide*, for a tutorial introduction to CombiGlide
- Frequently Asked Questions pages, at https://www.schrodinger.com/CombiGlide_FAQ.html

The manuals are also available in PDF format from the Schrödinger Support Center. Information on additions and corrections to the manuals is available from this web page.

If you have questions that are not answered from any of the above sources, contact Schrödinger using the information below.

E-mail: help@schrodinger.com
USPS: 101 SW Main Street, Suite 1300, Portland, OR 97204
Phone: (503) 299-1150
Fax: (503) 299-4532
WWW: http://www.schrodinger.com
FTP: ftp://ftp.schrodinger.com

Generally, e-mail correspondence is best because you can send machine output, if necessary. When sending e-mail messages, please include the following information, most of which can be obtained by entering `$SCHRODINGER/machid` at a command prompt:

- All relevant user input and machine output
- CombiGlide purchaser (company, research institution, or individual)
- Primary CombiGlide user
- Computer platform type
- Operating system with version number
- CombiGlide version number
- Maestro version number
- mmshare version number

Creating Custom Functional Groups and Custom Minimal Capping Groups

This appendix describes the procedures for adding custom functional groups or modifying existing functional groups.

The definitions of the functional groups are stored in two files in the CombiGlide distribution, `reagentprep.ini`, which is a plain text file, and `reagentprep.mae`, which is a structure file in Maestro format. These files may be found in the directory `$SCHRODINGER/combiglide-v`*version*`/data`. The functional group definitions include the following:

- **A long name.** This name is used to identify the functional group, and is used for its title. The long name is displayed in a tooltip when you pause the pointer over the reagent button in the Reagent Preparation panel. The name can be chosen any way you wish, but must not contain blanks. The long names for the supplied functional groups are constructed by adding the atom that is kept and the atom that is lost to the functional group name, separated by underscores, for example `Acid_Chloride_C_C`.
- **A short name.** This is the name that is displayed on the reagent button in the Reagent Preparation panel. The short names for the supplied functional groups are constructed by adding the atom that is kept and the atom that is lost to the shortened functional group name, separated by underscores, for example `Acid_Cl_C_C`. Because the reagent button displays the short name for custom functional groups, it is advisable to ensure that you can identify both the functional group and the bond that is broken in the name. Adopting the convention used for the supplied functional groups is one way to do this.
- **One or more SMARTS patterns that define the group.** The SMARTS patterns must be contiguous, without any spaces. At the top of `reagentprep.ini` is a description of the syntax with which you can use multiple SMARTS patterns to define the functional group.
- **A grow bond definition**. The grow bond defines the bond to be broken when the side chain from the reagent is added to the core. The bond is defined in terms of two atoms, the atom that is kept, and the atom that is deleted. This bond must be a single bond and must not be internal to a ring.
- **A structure that defines the minimal capping group.** This is the group that is added at the attachment position to define the minimally capped core, and in the construction of the singly substituted core molecules. The structure should consist of the desired functional group with the minimal capping group attached. For example, if you wanted to define an acid chloride with a methyl group as a minimal capping group, you would choose acetyl chloride as the structure. The structure should be 3D and minimized.

To add a new functional group, you must edit the file `reagentprep.ini` and add one or more lines that include the long name, the SMARTS pattern, and the grow bond definition, and you must add a structure to the file `reagentprep.mae`, with the long name, the short name, and the grow bond defined.

To change the capping group for an existing functional group, you must modify the structure in the file `reagentprep.mae`.

You can store the modified files in the directory `$HOME/.schrodinger/combiglide` or the directory from which you submit your reagent preparation job. When a reagent preparation job is run, CombiGlide looks first in the job submission directory for these files, then in `$HOME/.schrodinger/combiglide`, then in `$SCHRODINGER/combiglide-v`*version*`/data`. If you want to customize these files just for a particular job, you should store them in the job submission directory. If you want to use the custom functional groups for all your jobs, you should store the files in `$HOME/.schrodinger/combiglide`.

To add a functional group definition to reagentprep.ini:

1. Copy the file `reagentprep.ini` to the desired location.
2. Open `reagentprep.ini` in a text editor.
3. Insert a line into the file with the following format, followed by a blank line:

   `AddGN` *longname* *SMARTSpattern* *kept-atom*`:rpc1` *lost-atom*`:rpc2`

   An example for the acid chloride functional group is:

   ```
   AddGN Acid_Chloride_C_Cl [CX3,c](=O)(Cl)[#6] 1:rpc1 3:rpc2
   ```

   The variables that you need to replace have the following meanings:

longname	The long name of the functional group (see above).
SMARTSpattern	SMARTS pattern that defines the functional group.
kept-atom	The index of the atom in the grow bond that is kept when the bond is broken and the side chain is added to the core. The index is the index in the SMARTS string.
lost-atom	The index of the atom in the grow bond that is lost when the bond is broken and the side chain is added to the core. The index is the index in the SMARTS string.

   `rpc1` and `rpc2` are the *grow names* for the specified atoms, which are used to identify the kept atom and the lost atom when the side chain is "grown" onto the core.

4. Save and close the file.

To add a structure for the functional group to reagentprep.mae:

1. Copy the file `reagentprep.mae` to the desired location.
2. Start Maestro and open the Project Table panel (click the toolbar button shown below).

3. Add the desired structure to the Project Table.

   You can import the structure using the Import panel, or you can build it, or duplicate one of the structures in the Project Table and modify it. It might be useful to import the structures from `reagentprep.mae` to duplicate and modify.

   The structure must consist of the functional group and the minimal capping group only. The structure must be a 3D minimized structure. If it is not, you can convert and minimize it with LigPrep, or choose Clean Up Geometry from the Edit menu in the main window to minimize it with the UFF minimizer.

   If you build the structure, you must add it as an entry to the Project Table, by clicking the Create entry from Workspace toolbar button.

   The entry name does not matter, but the title must be the long name of the functional group. If you make the entry name the long name, the title is automatically copied from the long name and you do not need to change it in Step 5. You can also add the structure to the entry group that was imported, but it is not necessary to do so.
4. Enter the long name and the short name in the appropriate columns of the Project Table for the new structure.
5. Enter the long name as the title of the new structure.

   For both of these steps, you should ensure that the long name is the same as in the file `reagentprep.ini`.

Now you define the grow bond.

6. Ensure that the structure is displayed in the Workspace.
7. From the Display menu in the main window, choose Atom Labels.

   The Atom Labels panel opens.
8. In the Composition folder, select Grow Name and deselect all other labels.

9. Click All in the Label Atoms section.

   The atoms are now labeled with their grow names, if they have one.

10. Open the Build panel.

11. In the Atom Properties folder, choose Grow Name from the Property option menu.
12. Ensure that Atoms is chosen from the Pick option menu in the Apply grow name section.
13. Type `rpc1` into the Grow Name text box, and click on the atom in the Workspace that is to be kept when the bond is broken.
14. Type `rpc2` into the Grow Name text box, and click on the atom in the Workspace that is to be lost when the bond is broken.

    If you make a mistake with any of these assignments, just type the correct name in and pick the atom again. These two atoms must correspond to the atoms in the SMARTS pattern whose indices you entered in `reagentprep.ini`.

    To delete a grow name, enter two successive quotation marks "" (a null string) into the Grow Name text box.

Finally, you add the new structure to the file.

15. Choose Export Structures from the Project menu in the main window.

    The Export panel opens.

16. Navigate to your copy of `reagentprep.mae` and select it.
17. Under the File text box, select Append.
18. Under Structure sources to be exported, select Workspace.
19. Ensure that Export all entries to a single file is selected from the Files option menu.
20. Click Export.

    The entry displayed in the Workspace, which is your new structure, is appended to `reagentprep.mae`

If you want to add multiple functional groups, you can repeat the relevant steps above for other structures. When you come to adding the new structures, select the entries in the Project Table, and select Selected entries instead of Workspace under Structure sources to be exported in the Export panel (Step 18).

To modify a minimal capping group:

1. Copy the file `reagentprep.mae` to the desired location.
2. Start Maestro and open the Project Table panel (click the toolbar button shown below).

   If you start Maestro in the directory to which you copied `reagentprep.mae`, no navigation is required when you import or export the file.
3. Import `reagentprep.mae`.

   
4. Include the structure for the reagent you want to modify in the Workspace.
5. Modify the structure.

   You can use the Build panel to modify or add to the structure.
6. Minimize the structure.

   The preferred treatment is to use MacroModel with the OPLS_2005 force field. Minimization is important because the docking process does not change the bond lengths and bond angles in the single-position docking step.
7. Select all the structures that were imported.

   This should be all the structures in the Project Table, if you started with a new Maestro session.
8. Choose Export Structures from the Project menu in the main window.

   The Export panel opens.
9. Navigate to your copy of `reagentprep.mae` and select it.
10. Under the File text box, deselect Append.
11. Under Structure sources to be exported, select Selected entries.
12. Ensure that Export all entries to a single file is selected from the Files option menu.
13. Click Export.

    A dialog box is displayed to confirm the overwriting of `reagentprep.mae`
14. Click Yes.

The file `reagentprep.mae` is overwritten with the modified structure set.

To replace the existing structure for a capping group with another structure:

1. Follow the instructions for adding a structure above, from Step 1 through Step 14.

   When you enter the title, long name, and short name, use the values for the entry containing the structure you want to replace.

2. Delete the entry that you want to replace.

   The new entry has all the required information in it, and the order of the entries does not matter. If you want to move the new entry to the position that the deleted entry was in, select the entry and drag it to the desired location, or choose Move to Row from the Entry menu.

3. Select all the structures that were imported and the new structure.

   If you started with a new Maestro session, this should be all structures in the Project Table.

4. Follow the instructions for modifying a minimal capping group, from Step 8 through Step 14.

Copyright Notices

C and C++ Libraries for Parsing PDB Records

Glossary

Attachment—Combination of an attachment position and a reagent file.

Attachment position—Position on a core to which side chains are attached (synonymous with 'site of diversity').

Combinatorial library—A set of compounds prepared by combinatorial synthesis.

Combinatorial synthesis—The synthesis of sets of compounds utilizing the same reaction sequence where all possible combinations of the reagents are used.

Compound—In CombiGlide, a compound is explicitly defined by a combination of reagents that are attached to a core to form a product. A compound may encompass several structures. If a reagent exists in more than one form (different tautomers, conformers, stereoisomers, protonation state), the compound may also encompass more than one variant.

Core—The structural element of a combinatorial library that is constant throughout the library. It is the part of the core-containing molecule that remains when the side chains are removed.

Enumerate—To generate structures of all possible members of a combinatorial library.

Library selection—The results of filtering and selection that generates a list of reagents at each position for a focused library, including information on the filtering and selection strategy.

Fully substituted core—A core with side chains attached at all attachment positions.

Library definition—A collection of the core-containing structure and the reagent files that defines a combinatorial library. Can be used as input to library enumeration or combinatorial screening.

Minimal capping group—The smallest representative side chain for a particular functional group.

Minimally capped core—A core with the appropriate minimal capping group at each attachment position.

Reagent—The chemical reagent used in the combinatorial synthesis; a structure or set of structures with the same chemical composition and heavy-atom connectivity that is used to add a side chain to the core. A reagent may encompass more than one structure if it exists in more than one form (different tautomers, conformers, stereoisomers, protonation state).

Reagent variant—A single reagent structure, defined by a particular tautomer, conformer, stereoisomer, and protonation state of the reagent.

Side chains—The structural elements of a combinatorial library that are varied within the library.

Structure—A molecule with a particular geometry and stereochemistry, in a single tautomeric form, ionization state, and conformation.

Index

L

M

O

P

Q

R

S

T

U

W

120 West 45th Street
32nd Floor
New York, NY 10036

101 SW Main Street
Suite 1300
Portland, OR 97204

3655 Nobel Drive
Suite 430
San Diego, CA 92122

Dynamostraße 13
68165 Mannheim
Germany

QuatroHouse, Frimley Road
Camberley GU16 7ER
United Kingdom

SCHRÖDINGER.